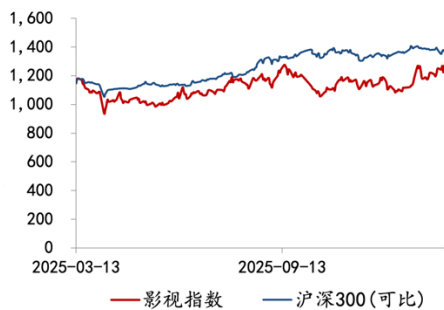


AI 视频：模型加速迭代，工具和 IP 价值凸显

■ 走势比较



■ 子行业评级

影视

看好

■ 相关研究报告

■ 证券分析师：郑磊

E-MAIL: zhenglei@tpyzq.com

执业资格证书编码：S1190523060001

■ 证券分析师：李林卉

E-MAIL: lilh@tpyzq.com

执业资格证书编码：S1190526020001

报告摘要

► 模型加速迭代，L3 短片级能力已能满足影视制作部分需求

模型方面，海外视频模型率先在物理模拟、保真性等前沿能力实现突破，国内模型持续追赶海外头部模型的同时，聚焦可控性、多模态交互、本土场景，形成差异化竞争。具体来看，目前视频模型原生分辨率最高达准 2K、单次生成时长最长达 25 秒。同时，支持音画同步输出，提升生成画面及叙事的可控性，更精准地模拟真实世界物理规律及人物动作与表情，并通过发布轻量模型及提质稳价间接降本。**内容生成方面**，主流模型目前已支持 L3 短片级内容创作，部分能力可达 L4 长片级内容要求，一定程度上已能满足影视行业制作需求。

► 影视行业迈入 AI 普及期，渗透率仍有较大提升空间

依托视频模型快速迭代，AI 视频工具逐步赋能影视制作：**1) 漫剧**：AI 在内容制作环节应用比例达 50%-80%，推动漫剧供给爆发式增长，目前 AI 漫剧数量占比超 70%。**2) 真人短剧**：由“AI+实拍”发展到全 AI 制作，AI 仿真人剧快速起量，头部作品《斩仙台》上线 6 天播放量迅速破亿。**3) 电影、电视剧**：仍以 AI 辅助制作为主。其中 AI 动画电影已率先落地，而真人电影仍处起步阶段。我们认为，AI 在影视素材拍摄环节显著降本，同时通过压缩内容制作周期、丰富内容题材及形式进行增效。**潜在空间来看**，全球视频制作市场规模约 3163 亿美元，目前全球 AI 视频市场规模约 45 亿美元，渗透率仅 1.4%，仍处较低水平。

► 视频工具为核心，IP 环节受益最大

AI 视频工具是模型能力转化为实际生产力的载体，因其既依赖于模型技术支持，又需专业视频内容制作实践积累，目前形成了**视频模型、IP 及影视、第三方工具公司**协同发展的格局。未来，拥有技术领先优势的 AI 视频工具公司更易依托自身创作能力与平台生态，产出优质视频内容。同时，IP 公司因具备海量可视频化的内容库资源，有望充分受益于 AI 视频工具的成熟。

► 核心观点

我们认为，2025 年以来海内外视频模型在性能上加速迭代突破，已满足 L3 短片级内容制作能力，推动全球影视行业迈入 AI 普及期。目前，AI 在影视行业渗透率仍处于个位数，随着模型和视频工具的进一步迭代，行业渗透率有望迎来爆发式增长。而 AI 视频工具作为载体，产业链核心价值凸显。同时，IP 公司有望充分受益于此轮浪潮，实现内容资产价值重估。

► 风险提示

AI 技术发展不及预期、内容监管趋严、行业竞争加剧的风险。

目录

1. 视频模型：海内外差异化竞争，已支持短片级内容创作	5
1.1 海外：率先实现物理模拟、保真度等前沿突破，VEO 3 领跑全球	6
1.2 国内：差异化突围，强化可控性、多模态交互与本土适配	11
1.3 模型已支持 L3 短片级内容创作，仍处技术快速迭代期	17
2. 影视场景：细分领域渗透率提升，降本增效驱动增长	23
2.1 影视细分领域：AI 漫剧率先落地，AI 仿真人剧快速起量	23
2.2 AI 赋能降本增效，渗透率仍有较大提升空间	28
3. AI 视频标的梳理：视频工具为核心，IP 环节受益最大	35
3.1 视频模型公司：占据技术和商业化场景优势，处主导地位	35
3.2 IP 及影视公司：将一线需求融入工具迭代，赋能 AI 影视工业化	37
3.3 第三方工具公司：初创公司为主，聚焦垂类应用场景	39
4. 风险提示	41

图表目录

图表 1: AI 视频模型发展历程.....	5
图表 2: 海内外主流 AI 视频模型.....	6
图表 3: OPENAI SORA 各代版本信息梳理.....	6
图表 4: OPENAI SORA 各代版本示例.....	7
图表 5: 谷歌 VEO 各代版本信息梳理.....	7
图表 6: 谷歌 VEO 各代版本示例.....	8
图表 7: RUNAWAY GEN 各代版本信息梳理.....	9
图表 8: RUNAWAY GEN 各代版本示例.....	9
图表 9: LUMA RAY 各代版本信息梳理.....	10
图表 10: LUMA RAY 各代版本示例.....	10
图表 11: 快手可灵各代版本信息梳理.....	11
图表 12: 快手可灵各代版本示例.....	12
图表 13: 字节 SEEDANCE 各代版本信息梳理.....	13
图表 14: 字节 SEEDANCE 各代版本示例.....	14
图表 15: 阿里万相各代版本信息梳理.....	15
图表 16: 阿里万相各代版本示例.....	15
图表 17: MINIMAX 海螺各代版本信息梳理.....	16
图表 18: MINIMAX 海螺各代版本示例.....	17
图表 19: 海外视频模型优先突破分辨率, 国内优先拓展时长.....	18
图表 20: 主流视频模型生成速度及成本对比.....	19
图表 21: 目前主流视频模型均已支持多模态输入、音画同步输出.....	20
图表 22: 可灵 3.0 OMNI 自定义分镜功能.....	20
图表 23: 海螺 1.0-LIVE 模型.....	20
图表 24: SORA 和 SORA2 物理规律合规率对比.....	21
图表 25: SORA 和 SORA2 生成滑板运动对比.....	21
图表 26: 海螺 2.3 生成复杂动作、人物表情.....	21
图表 27: AIGC 内容创作分级标准.....	22
图表 28: AI 漫剧成熟实践 workflow.....	23
图表 29: 2025 年 7-12 月漫剧供给数量.....	24
图表 30: 主流漫剧分类.....	24
图表 31: 《斩仙台下, 我震惊了诸神!》.....	25
图表 32: 2025. 07-2025. 12 抖音各类型短剧部均播放量.....	25
图表 33: AI 仿真人剧代表作品.....	26
图表 34: AI+短剧和 AI+电影、电视剧对比.....	27
图表 35: 《传说》27 岁成龙形象.....	27
图表 36: 《EL ETERNAUTA》AI 画面.....	27
图表 37: 中国首部全流程 AIGC 动画电影《团圆令》.....	28
图表 38: 中国首部全 AI 院线动画电影《山海经·玄武》.....	28
图表 39: 全球首部 AI 长篇电影《OUR T2 REMAKE》.....	28
图表 40: 全球首部院线 AIGC 电影《海上女王郑一嫂》.....	28
图表 41: AI 赋能视频制作全流程降本.....	29
图表 42: 短剧成本构成及占比.....	29
图表 43: 传统动漫短剧和 AI 漫剧制作成本、团队人数对比.....	30

图表 44: 传统及 AI 影视制作流程对比	31
图表 45: 传统动漫短剧和 AI 漫剧制作效率、周期对比	31
图表 46: 纯人工和人工+AI 制作漫剧效能对比	32
图表 47: 网文 IP 内容题材结构	33
图表 48: 多样化内容形式引入	33
图表 49: AI 转绘漫剧《心动的信号》	33
图表 50: 2026 全球电影及视频制作市场 3163 亿美元	34
图表 51: 2023-33 全球 AI 视频市场规模(十亿美元)	34
图表 52: AI 漫剧和 AI 真人短剧对比	35
图表 53: AI 视频工具赛道三类公司	35
图表 54: 可灵 AI 工作台	36
图表 55: 全球首部 AI 单元剧集《新世界加载中》	36
图表 56: 即梦 AI 工作台	37
图表 57: 《金鱼》	37
图表 58: 次元神笔工作台	38
图表 59: AI 3D 动漫《仙尊归来: 只手镇人间》	38
图表 60: 博乐工作台	39
图表 61: 博乐创作作品部分海报展示	39
图表 62: 新境 AI 工作台	40
图表 63: AI 科幻短剧《浮光》	40
图表 64: 可梦 AI 工作台	40
图表 65: 可梦 AI 精品 AI 短剧	40
图表 66: 灵境 AI 工作台	41
图表 67: AI 转绘漫剧《武极天尊》转绘效果	41

1. 视频模型：海内外差异化竞争，已支持短片级内容创作

DiT 成为主流架构，目前 AI 视频模型处于快速迭代期。2022 年底以 GPT 为代表的大语言模型实现关键突破，开启 AI 发展浪潮。在此背景下，AI 视频模型也迎来了快速发展，其发展历程可分为三大阶段：**1) 技术架构扩散期：**2014-2022 年，视频模型先后历经 GAN、Transformer、Diffusion 等技术架构迭代。2023 年 6 月，Runway 基于 Diffusion 架构推出 Gen-2，突破文生、图生视频能力，成为全球首个实现商业化的文生视频模型。**2) DiT 架构普及期：**2024 年 2 月，OpenAI Sora 将 DiT 架构引入视频生成领域，验证了技术的有效性。此后，基于 DiT 架构，各大模型厂商相继发布 AI 视频模型，如海外谷歌 Veo，国内快手可灵 1.0、MiniMax 海螺 1.0 等，奠定了 DiT 架构的主流地位。**3) 技术快速迭代期：**2025 年以来，模型更新迭代速度显著加快，并实现统一多模态、音画同出、多镜头叙事等核心突破，模型生成可控性、美学风格、物理模拟等逐步提升。

图表 1：AI 视频模型发展历程



资料来源：太平洋证券研究院整理

参考 Artificial Analysis 视频模型排行榜，以及网页访问量数据，选取八个 AI 视频模型，并按海外和国内模型进行划分。通过梳理各个模型的版本迭代细节，总结目前主流 AI 视频模型的演进趋势，以及内容生成能力所处的阶段。

图表 2：海内外主流 AI 视频模型

	公司名称	最新模型名称	发布时间	1 月访问量 (万)
海外	OpenAI	Sora2	2025.09.30	2967
	Google	Veo3.1	2025.10.15	-
	Runway	Gen-4.5	2025.12	582
	Luma	Ray3.14	2026.01.26	247
国内	快手	可灵 3.0	2026.02.04	1548
	字节跳动	Seedance 2.0	2026.02.12	-
	阿里巴巴	万相 2.6	2025.12.16	381
	海螺 AI	海螺 2.3	2025.10.28	531

资料来源：各公司官网，AI 产品榜，太平洋证券研究院

1.1 海外：率先实现物理模拟、保真度等前沿突破，Vevo 3 领跑全球




OpenAI Sora：确立 DiT 架构主流地位，上线应用开启 C 端尝试。 1) 开启 AI 视频新纪元：2024 年 2 月，OpenAI 发布 Sora 预览版，将 DiT 架构引入视频生成领域，奠定视频模型发展基础。2) 已完成两个版本更新：2024 年 12 月，Sora Turbo 正式向用户开放使用，实现从预览版到商业化落地。时隔 10 个月，Sora2 发布，单次生成时长 25 秒及原生生成分辨率准 2K，均为视频模型最高水平。同时，在物理规律和复杂指令遵循上实现显著提升。从生成样片来看，角色复杂动作、环境或物体表面光影等生成内容都较前代模型更接近真实世界。3) 布局 C 端应用：Sora2 推出的同时，面向 C 端用户推出 Sora 应用。但因应用内大多为相似的 AI 视频导致用户审美疲劳、产品定位模糊等原因，Sora 应用近期刊下载量、活跃用户等数据持续下滑。

图表 3：OpenAI Sora 各代版本信息梳理

模型版本	发布时间	生成时长	分辨率	提示词类型	音频	定价
Sora 预览版	2024.02.15	-	最高 1080p	文本	无	-
Sora Turbo	2024.12.09	5/10/15/20 秒	480p、720p、1080p	文本、图片、视频	无	-
Sora2	2025.09.30	10/15/25 秒	720p、1080p、准 2K	文本、图片	原生音频	1) Sora2(720p)：0.1 美元/秒 2) Sora2 Pro(720p)：0.3 美元/秒 3) Sora2 Pro(准 2K)：0.5 美元/秒

资料来源：OpenAI 官网，太平洋证券研究院

图表 4: OpenAI Sora 各代版本示例

模型版本	特点	新增功能	示例
Sora 预览版	1) 高真实感 2) 强调世界模拟	首次引入文本生成视频	
Sora Turbo	1) 更快的生成速度 (较预览版提高约 40%) 2) 更强大的功能 3) 支持多分辨率、比例选择	1) 画面延展功能 2) 视频循环功能 3) 视频元素编辑功能 4) 故事板功能 (在时间轴设置场景、角色、动作序列) 5) 视频合并功能 (以过渡、点缀等方式合并两个视频)	
Sora2	1) 同步音频生成 2) 更好的遵循物理规律 3) 更强的可控性、指令遵循	1) 音视频同步生成功能 2) 角色扮演功能 (让用户成为生成视频内的角色) 3) 发布 Sora App 4) 发布增强实验版模型 Sora2 Pro	

资料来源: OpenAI 官网, 太平洋证券研究院





谷歌 Veo: 专注电影级创作控制, 生成视频质量领跑全球。2024 年 5 月, 谷歌发布 Veo 预览版, 模型专注于电影级创作控制。此后, Veo 围绕这一核心定位完成三次更新迭代。在模型能力上, 依托对电影术语的深度理解, 持续优化物理真实感、镜头语言表现力、叙事连贯性等。其中, 2025 年 5 月发布的 Veo3 是最早实现音画同出的视频模型, 且分辨率可通过超分功能升至 4K; 最新迭代版 Veo3.1 位列 LMArena 文生视频榜榜首, 生成视频质量领跑全球。在创意控制上, 推出时长延长、对象增减、镜头控制等功能, 使模型逐步向专业视频编辑工具演进。在产品生态上, 发布面向电影制作人的 AI 电影制作工具 Flow, 集成 Gemini、Imagen、Veo, 提供剧本创作、角色创建、视频生成与编辑等功能。并搭建 Flow TV, 提供展示和浏览生成视频的渠道。截至 2025 年 10 月, Flow 累计生成视频超 2.75 亿个。

图表 5: 谷歌 Veo 各代版本信息梳理

模型版本	发布时间	生成时长	分辨率	提示词类型	音频	定价
Ve0 预览版	2024.05.14	-	最高 1080p	文本	无	-
Ve02	2024.12.16	5-8 秒	720p	文本、图片	无	0.35 美元/秒
Ve03	2025.05.21	4/6/8 秒	720p、1080p (仅限 16:9)、4K	文本、图片	原生音频	0.4 美元/秒
Ve03.1	2025.10.15	4/6/8 秒	720p、1080p (仅限 8 秒)、4K	文本、图片、视频	原生音频	1) 720p 和 1080p: 0.4 美元/秒 2) 4K: 0.6 美元/秒

资料来源: 谷歌官网, 太平洋证券研究院

图表 6: 谷歌 Veo 各代版本示例

模型版本	特点	新增功能	示例
Veo 预览版	<ol style="list-style-type: none"> 1) 理解自然语言和视觉语义 2) 理解电影术语 3) 支持多种视觉风格 	-	
Veo2	<ol style="list-style-type: none"> 1) 提升电影理解 2) 更好的遵循物理规律、真实感 3) 更低的幻觉 4) 更高的分辨率 	<ol style="list-style-type: none"> 1) 生成视频带有 SynthID 水印 2) 延长功能 (可让生成视频时长延长至几分钟) 	
Veo3	<ol style="list-style-type: none"> 1) 首次支持原生音频, 如对话、音效、环境声等生成 2) 更强的提示词遵循和叙事控制 3) 更好的遵循物理规律、真实感 4) 更优的角色一致性 	<ol style="list-style-type: none"> 1) 图片参考功能 2) 镜头控制功能 (精准定义镜头动线, 如旋转、推轨、变焦等) 3) 拓展功能 (调整视频比例, 如从竖屏切换到横屏) 4) 可直接生成竖屏视频, 适配社交媒体视频比例 5) 超分功能 (通过超分技术将分辨率提升至 4K) 6) 发布 AI 电影制作工具 Flow (整合 Veo、Imagen、Gemini 模型, 无缝创建电影片段和场景) 	
Veo3.1	<ol style="list-style-type: none"> 1) 更精准的同步生成更丰富、高质量的音频 2) 更强的提示词遵循和叙事控制 3) 更好的遵循物理规律、真实感 4) 更优的角色、环境一致性 	<ol style="list-style-type: none"> 1) 多图参考功能 (支持输入最多三张参考图片) 2) 增减对象功能 (添加/移除/更改视频中的背景图像或物体) 3) 首尾帧连接功能 4) 发布 Veo 3.1 Fast 	

资料来源: 谷歌官网, 太平洋证券研究院






Runaway Gen: 强调视频编辑能力, Gen2 为全球首个商业化文生视频模型。Runaway 从视频编辑工具起家, 2023 年 2 月发布 Gen-1, 实现视频到视频的风格化转换, 支持将提示词描述的风格应用于源视频。此后, Gen 完成四次更新迭代, 始终强调视频编辑能力, 于 2023 年 6 月最早推出镜头控制、动作笔刷等功能。未来, Gen 计划将已有的编辑控制功能拓展至图生视频和视频生视频领域, 持续强化视频编辑优势。此外, 2023 年 6 月发布的 Gen-2 是全球首个实现商业化的文生视频模型, 其基于 Diffusion 架构突破文生、图生视频能力, 并推出订阅制+积分制的混合付费模式。

图表 7: Runaway Gen 各代版本信息梳理

模型版本	发布时间	生成时长	分辨率	提示词类型	音频	定价*
Gen-1	2023.02	4 秒	720p	视频	无	-
Gen-2	2023.06	4 秒	720p	文本、图片、视频	无	-
Gen-3 Alpha	2024.06	5 秒	720p	文本、图片、视频	无	1) Gen-3: 0.20 美元/秒 2) Gen-3 Turbo: 0.10 美元/秒
Gen-4	2025.03	5/10 秒	720p	文本、图片 (重点)	无	1) Gen-4: 0.24 美元/秒 2) Gen-4 Turbo: 0.10 美元/秒
Gen-4.5	2025.12	5/10 秒	720p	文本 (重点)、图片	无	0.50 美元/秒

资料来源: Runaway 官网, 太平洋证券研究院
 *以\$0.02 美元/信用消耗计算

图表 8: Runaway Gen 各代版本示例

模型版本	特点	新增功能	示例
Gen-1	1) 主要应用于视频到视频的风格转换	-	
Gen-2	1) 首次实现文本、图片生成视频 2) 丰富了编辑控制功能	1) 超分功能 (通过超分技术将分辨率提升至 4K) 2) 延长/拼接功能 (延长生成视频的时长) 3) 比例自定义功能 (根据选择, 生成多比例视频) 4) 镜头控制功能 (控制镜头移动、速度、方向等) 5) 动作笔刷功能 (在特定区域绘制对象运动方向)	
Gen-3 Alpha	1) 显著提升一致性、保真度、运动真实感 2) 更好的遵循物理规律	1) 高级镜头控制功能 (增加镜头控制的运动方向、可在时间轴定义对象的出现、变化、运动) 2) 延长/拼接功能 (延长次数增加, 视频时长更长) 3) 发布 Gen-3 Turbo (提升生成速度、降低成本)	
Gen-4	1) 提升角色、环境的一致性, 跨镜头的连贯性 2) 实现电影级的运动质量 3) 更好的遵循物理规律 4) 更强的提示词遵循和叙事能力	1) 图片参考功能 2) 延长/拼接功能 (延长次数增加, 视频时长更长) 3) 发布 Gen-4 Turbo (提升生成速度、降低成本)	
Gen-4.5	1) 提升一致性、保真度、风格控制、编辑控制能力 2) 更好的遵循物理规律 3) 更强的提示词遵循	1) 计划推出原生音频生成功能 2) 计划将已有的编辑控制功能拓展至图生视频、视频生视频	

资料来源: Runaway 官网, 太平洋证券研究院

Luma Ray: 构建垂类场景优势, Ray3 为全球首个原生 HDR 及推理驱动视频模型。
 Ray 聚焦专业影视创作场景, 构建起垂类场景优势。在模型能力上, Ray3 是全球首个支持原生 HDR 的模型, 且分辨率可通过超分功能升至 4K, 生成视频质量达到部分影视级制作标准; 同时, Ray3 也是全球首个推理驱动模型, 区别于非推理模型由提示词

直接映射至像素生成，其通过提示词理解、意图推理、输出规划，实现生成视频的保真度 SOTA，更适配影视、CG 等高精度内容创作需求。在产品生态上，2024 年 11 月，Dream Machine v1 升级为 AI 多模态创作平台，完成 ios 应用开发、界面优化、Photon 图像生成模型集成等迭代。平台依托多个自研多模态模型与多样化专业编辑功能，为创作者提供从生成到编辑的全流程工具。




图表 9: Luma Ray 各代版本信息梳理

模型版本	发布时间	生成时长	分辨率	提示词类型	音频	定价*
Dream Machine v1	2024. 06. 12	5 秒	720p	文本、图片	无	-
Ray2	2025. 01. 15	5/10 秒	最高 4K	文本、图片、视频	无	1) 540p: 0.08 美元/秒 2) 720p: 0.14 美元/秒 3) 1080p: 0.17 美元/秒 4) 1080p 超分至 4K: 0.02 美元/秒
Ray3	2025. 09. 18	5/10 秒	最高 4K	文本、图片、视频	无	-
Ray3. 14	2026. 01. 26	5/10 秒	最高 4K	文本、图片、视频	无	-

资料来源: Luma 官网, 太平洋证券研究院

*生成 16:9 视频的定价

图表 10: Luma Ray 各代版本示例

模型版本	特点	新增功能	示例
Dream Machine v1	1) 生成速度快, 20-60 秒可完成生成 2) 简单易用, 创意自由度高	-	-
Ray2	1) 算力资源是一代模型的 10 倍, 实现更快的生成速度 2) 强调运动连贯性、细节逼真度、时序逻辑性 3) 更强的提示词遵循 4) 更高的分辨率	1) 超分功能 (通过超分技术将分辨率提升 4K) 2) 延长/拼接功能 (视频时长延长至 1 分钟) 3) 镜头控制功能 4) 发布 Ray Flash 2 (生成速度提升 3 倍、成本下降 67%)	
Ray3	1) 支持原生 HDR 视频生成 (可通过输入文本、SDR 图片、SDR 视频实现), 并输出可供专业创作者使用的 EXR 原始素材格式 2) 全球首个推理驱动视频模型 3) 实现保真度的 SOTA (提升真实度、物理规律遵循、一致性)	1) 草稿模式 (先生成预览, 挑选后再生成高质量成片; 生成速度提升 5 倍、成本下降 80%) 2) 视频编辑功能 (可修改视频背景、环境、对象等; 基于参考视频生成特定风格的视频) 3) 首尾帧连接功能 4) 图片参考功能	
Ray3. 14	1) 生成速度提升 4 倍, 成本下降 67% 2) 全流程支持原生 1080p 生成 3) 更强的提示词遵循、一致性	-	

资料来源: Luma 官网, 太平洋证券研究院

1.2 国内：差异化突围，强化可控性、多模态交互与本土适配

快手可灵：深耕多模态及可控生成，主体一致性超96%。2024年6月，快手发布基于DiT架构的可灵1.0，是全球首个用户可用的DiT架构视频模型。此后，可灵以**2-4个月为周期进行迭代**，至今已完成10个版本的发布，其深耕多模态及可控生成，推动一致性持续提升。**在多模态方面**，可灵2.0提出全新的多模态视觉语言交互范式MVL，该范式将输入内容从自然语言升级为多模态视觉语言，弥补了纯文本提示词的信息短板，实现对角色、场景、镜头等内容的精准约束。**在可控生成方面**，模型相继推出多模态编辑、主体库、动作控制、自定义分镜等功能，通过对画面内容、角色特征、动作姿态、镜头语言等内容的精准调控，有效降低视频生成的随机性。依托多模态及可控生成能力，可灵模型主体一致性表现突出，其中可灵o1主体一致性超96%。

图表 11：快手可灵各代版本信息梳理

模型版本	发布时间	生成时长	分辨率	提示词类型	音频	定价*
可灵 1.0	2024.06.06	5/10 秒	720p、1080p	文本、图片	无	0.2 元/秒
可灵 1.5	2024.09.20	5/10 秒	720p、1080p	文本、图片	无	0.4 元/秒
可灵 1.6	2024.12.19	5/10 秒	720p、1080p	文本、图片	无	0.4 元/秒
可灵 2.0	2025.04.15	5/10 秒	720p、1080p	文本、图片	无	-
可灵 2.1	2025.05.29	5/10 秒	720p、1080p	图片	无	0.4 元/秒
可灵 2.5 Turbo	2025.09.26	5/10 秒	720p、1080p	文本、图片	无	0.3 元/秒
可灵 o1	2025.12.01	3-10 秒	720p、1080p	文本、图片	无	1) 无参考：0.6 元/秒 2) 有参考：0.9 元/秒
可灵 2.6	2025.12.03	5/10 秒	720p、1080p	文本、图片	原生音频	0.3 元/秒
可灵 3.0	2026.02.04	3-15 秒	720p、1080p	文本、图片、视频	原生音频	1) 文生无声：0.6 元/秒 2) 图生有声：0.9 元/秒
可灵 3.0 Omni	2026.02.04	3-15 秒	720p、1080p	文本、图片、音频、视频	原生音频	1) 无声：0.6 元/秒 2) 有声：0.8 元/秒

资料来源：可灵官网，太平洋证券研究院

*均为标准模式的定价

图表 12：快手可灵各代版本示例

模型版本	特点	新增功能	示例
可灵 1.0	<ul style="list-style-type: none"> 1) 全球首个用户可用的 DiT 视频模型 2) 中文理解准确率达 93% 	<ul style="list-style-type: none"> 1) 视频续写功能 (可通过微调提示词进行视频续写创作; 可续写 4-5 秒, 且支持多次续写, 最长可续写至 3 分钟) 	
可灵 1.5	<ul style="list-style-type: none"> 1) 显著提升画面质量与人物美感 2) 更合理的动作表现 3) 更精准的提示词和指令遵循能力 	<ul style="list-style-type: none"> 1) 尾帧生成功能 (据尾帧图片生成 5、10 秒视频) 2) 运动笔刷功能 (为图片的对象绘制运动轨迹) 3) 对口型功能 (人物语音与口型精准匹配) 4) 运镜控制功能 (支持水平、垂直等 6 种运镜) 5) 人脸模型 (保持人脸的一致性) 	
可灵 1.6	<ul style="list-style-type: none"> 1) 在内部测评中, 图生视频较 1.5 模型整体效果提升 195% 2) 显著提升画面质量、动态质量 3) 更合理的动作表现 4) 更精准的提示词和指令遵循能力 	<ul style="list-style-type: none"> 1) 多图参考功能 (支持输入最多四张参考图片) 2) AI 模特功能 (生成 AI 模特, 并通过图生视频生成动态服装展示视频) 3) 强化运动笔刷和对口型功能的精度 	
可灵 2.0	<ul style="list-style-type: none"> 1) 在画面质量、指令遵循、动态质量保持全球领先 2) 发布全新的生成交互范式 MVL/多模态视觉语言 (交互方式从文本升级为视觉语言, 提高传达指令精确性) 	<ul style="list-style-type: none"> 1) 多模态的视频编辑功能 (通过输入文字或图片, 实现生成视频内的元素增加、删减、替换, 实现二次编辑和处理) 	
可灵 2.1	<ul style="list-style-type: none"> 1) 提升画面质量和视觉效果 2) 更合理的动作表现 3) 更精准的提示词和指令遵循能力 	<ul style="list-style-type: none"> 1) 首尾帧连接功能 2) 发布创意工作台灵动画布 1.0, 实现生成、编辑、多人协同等功能的全流程整合, 支持无限画布拓展和多镜头时序排布 	
可灵 2.5 Turbo	<ul style="list-style-type: none"> 1) 更精确的提示词理解、时序控制 2) 更流畅稳定的高速动态质量 3) 提升在艺术风格上的一致和连贯性 	-	
可灵 o1	<ul style="list-style-type: none"> 1) 全球首个统一多模态视频模型, 在单一输入框内可无缝融合多种模态信息和任务 2) 结合思维链技术, 赋予模型强大的常识推理与事件推演能力 	<ul style="list-style-type: none"> 1) 主体库功能 (赋予模型长期记忆, 实现跨场景角色一致性) 	
可灵 2.6	<ul style="list-style-type: none"> 1) 实现音画同出 (支持单人对白、旁白解说、音乐表演等音视频同出) 2) 具备影视级的叙事张力 (动作自然、人物情绪表情细腻) 	<ul style="list-style-type: none"> 1) 音色控制功能 (支持自定义角色声线并保持一致, 多角色下角色音色的识别和绑定) 2) 动作控制功能 (实现 30 秒武打级动作、表情、手势控制) 3) 发布灵动画布 Agent, 包括多轮对话智能编辑、电商自动组图等模块 	
可灵 3.0	<ul style="list-style-type: none"> 1) 采用统一多模态训练框架, 支持文本、图像、视频片段等多种输入形式 2) 升级音画同步生成能力, 支持多种语言及方言的精准口型匹配 3) 提升视频画质和细节, 如文字生成清晰度、动作精细度、光影效果等 4) 强化主体一致性控制 	<ul style="list-style-type: none"> 1) 智能分镜系统 (自动解析文本叙事逻辑, 智能调度特写、全景等景别, 配合动态机位调整, 输出具备专业分镜脚本的连续画面) 2) 多图+视频参考功能 	
可灵 3.0 Omni	<ul style="list-style-type: none"> 1) 专属 IP 级内容创作引擎 2) 强化文本、图片、音频、视频全模态交互 	<ul style="list-style-type: none"> 1) 创建视频主体特征库功能 (支持提取 3-8 秒视频片段中的角色特征, 建立专属数字资产库, 后续生成可直接调用特征库, 并保持主体一致性) 2) 原生自定义分镜功能 	

资料来源: 可灵官网, 太平洋证券研究院

字节 Seedance: 具备原生导演级多镜头生成能力, 生成可用率达到业界 SOTA。1) 多镜头能力: 2025 年 5 月, 字节发布 Seedance 1.0 Pro, 通过将文本提示拆解为专业镜头序列、跨镜头保持一致性等技术, 实现 2-3 个连贯镜头的叙事生成。9 个月后发布的 Seedance 2.0 进一步提升多镜头叙事能力, 通过分析叙事逻辑、自分镜和自运镜, 实现原生导演级多分镜无缝生成, 输出具有景别变化、机位运动、时空连贯的多镜头视频。**2) 全模态参考能力:** Seedance 2.0 具备全模态参考功能, 支持输入文本、以及最多 9 张图片、3 段视频、3 段音频作为生成参考, 实现视频生成的精准可控。依托模型两大核心能力, Seedance 2.0 的视频生成可用率大幅提升, 根据极客公园测算, 其生成 15 秒视频的可用率或达 90%, 远超此前业内均值 20%。

图表 13: 字节 Seedance 各代版本信息梳理

模型版本	发布时间	生成时长	分辨率	提示词类型	音频	定价*
Seedance 1.0 Lite	2025. 04. 28	5 秒	720p	文本、图片	无	0.22 元/秒
Seedance 1.0 Pro	2025. 05. 28	2-12 秒	480p、720p、1080p	文本、图片	无	480p: 0.14 元/秒; 720p: 0.32 元/秒; 1080p: 0.73 元/秒
Seedance 1.0 Pro Fast	2025. 10. 15	2-12 秒	480p、720p、1080p	文本、图片	无	480p: 0.04 元/秒; 720p: 0.09 元/秒; 1080p: 0.20 元/秒
Seedance 1.5 Pro	2025. 12. 16	4-12 秒	480p、720p、1080p	文本、图片	原生音频	1) 无声: 480p: 0.08 元/秒; 720p: 0.17 元/秒; 1080p: 0.39 元/秒 2) 有声: 480p: 0.16 元/秒; 720p: 0.35 元/秒; 1080p: 0.78 元/秒
Seedance 2.0	2026. 02. 12	4-15 秒	480p、720p	文本、图片、音频、视频	原生音频	-

资料来源: 火山引擎官网, 太平洋证券研究院

*根据每百万 tokens 定价, 以及每秒视频消耗 token 进行计算

图表 14: 字节 Seedance 各代版本示例

模型版本	特点	新增功能	示例
Seedance 1.0 Lite	<ul style="list-style-type: none"> 1) 生成速度快, 20-60 秒可完成视频生成 2) 简单易用 	<ul style="list-style-type: none"> 1) 首尾帧连接功能 	
Seedance 1.0 Pro	<ul style="list-style-type: none"> 1) 2-3 个连贯镜头的叙事能力, 且在镜头切换中保持一致性 2) 优化运动生成的稳定性, 使生成视频具备物理真实感 3) 更精准的语义理解与指令遵循能力 4) 更多样的风格化表达 	<ul style="list-style-type: none"> 1) 升级首尾帧连接功能 (改善人脸一致性、提升运动自然合理度、优化视频节奏) 2) 上线 Agent 模式, 支持一句话拆解需求、自动执行全流程 	
Seedance 1.0 Pro Fast	<ul style="list-style-type: none"> 1) 在视频生成质量、速度、成本之间取得平衡; 继承 1.0 pro 模型核心优势的基础上, 生成速度最高提升约 3 倍, 生成价格降低约 72% 2) 进一步提升人物面部、发丝纹理等细节表现力 	-	
Seedance 1.5 Pro	<ul style="list-style-type: none"> 1) 实现音画高精度同步生成 2) 具备影视级的运镜控制和动态张力 (动作自然、人物情绪表情细腻) 3) 更强的语义理解和叙事协调性 	<ul style="list-style-type: none"> 1) 音频同步生成功能 (覆盖环境音、动作音、乐器音、背景音、人声等, 支持多人、多语言、方言对白) 2) 计划上线样片功能, 实现效率提升 65%、成本减少 60% 	
Seedance 2.0	<ul style="list-style-type: none"> 1) 具备原生多镜头叙事能力, 可实现自分镜和自运镜 (通过自动分析叙事逻辑, 生成具有景别变化、机位运动、时空连贯的多镜头序列) 2) 显著增强物理准确度、逼真度、可控性 3) 显著提升指令遵循与一致性表现 4) 生成可用率达到业界 SOTA 水平 	<ul style="list-style-type: none"> 1) 全模态参考功能 (支持输入最多 9 张图片、3 段视频、3 段音频) 2) 视频编辑功能 (支持主体替换、对象增删改、局部重绘或修复等) 3) 视频延长功能 	

资料来源: 火山引擎官网, 太平洋证券研究院

阿里万相: 以中文创作需求为核心, 最早实现音画同出、单次生成时长 15 秒。

2024 年 9 月, 阿里发布万相视频模型, 该模型以中文创作需求为核心, 针对中式元素的概念理解和生成表现力做了特别优化, 擅长生成国风视频。时隔五个月后发布的万相 2.1, 延续了初代模型关注中文创作的特点, 为首个支持中文文字及中英文文字特效生成的视频模型。此后, 万相完成三次更新迭代, 实现多个国内领先的技术突破:




- 1) 万相 2.2 是业界首个使用 MoE 架构的视频模型, 在同参数规模下, 实现算力消耗节省约 50%;
- 2) 万相 2.5 采用原生多模态架构, 是国内首个实现音画同步生成的模型;
- 3) 万相 2.6 是国内首个实现单次生成视频时长 15 秒的模型。

图表 15: 阿里万相各代版本信息梳理

模型版本	发布时间	生成时长	分辨率	开/闭源	提示词类型	音频	定价
万相视频模型	2024.09.19	5 秒	720p	-	文本、图片	基础音效	免费
万相 2.1	2025.02.25	3/4/5 秒	480p、720p	开源	文本、图片、视频	基础音效与背景音乐	0.7 元/秒
万相 2.2	2025.07.28	5 秒	480p、720p、1080p	开源	文本、图片	基础音效与背景音乐	480p: 0.2 元/秒; 720p: 0.4 元/秒; 1080p: 0.7 元/秒
万相 2.5	2025.09.24	5/10 秒	480p、720p、1080p	闭源	文本、图片、音频	原生音频	480p: 0.3 元/秒; 720p: 0.6 元/秒; 1080p: 1 元/秒
万相 2.6	2025.12.16	2/5/10/15 秒	720p、1080p	闭源	文本、图片、音频	原生音频	1) 文生/参考生视频: 720p: 0.6 元/秒; 1080p: 1 元/秒 2) 图生视频无声: 720p: 0.15 元/秒; 1080p: 0.25 元/秒 3) 图生视频有声: 720p: 0.3 元/秒; 1080p: 0.5 元/秒

资料来源: 阿里云官网, 太平洋证券研究院

图表 16: 阿里万相各代版本示例

模型版本	特点	新增功能	示例
万相视频模型	<ol style="list-style-type: none"> 1) 具备复杂语义理解能力, 并针对中式元素的概念理解与生成表现力做特别优化, 可轻松创作国风视频 2) 实现大幅度主体运动和运镜控制, 还原物理规律 3) 生成与视觉内容高度匹配的声音特效, 实现音画同步 	<ol style="list-style-type: none"> 1) 灵感扩写功能 (自动扩写简单提示词, 大幅提升生成效果) 2) 首帧延续生成功能 (根据首帧图片延续生成视频, 生成更可控, 画面更精准) 	
万相 2.1	<ol style="list-style-type: none"> 1) 包含文生视频、图生视频、视频生成与编辑共五款模型 2) 首个支持中文文字、中英文文字特效生成的视频模型 3) 视频生成与编辑模型具备文生视频、图像参考视频, 视频重绘、局部编辑、背景延展、时长延展等生成和编辑能力 4) 1.3B 参数模型能在消费级显卡运行 	<ol style="list-style-type: none"> 1) 图片参考功能 2) 重绘功能 (支持基于人体姿态、运动光流、画面景深、运动轨迹、着色等控制生成) 3) 局部编辑功能 (视频元素的替换、增加和删除等) 4) 视频延展功能 (在空间、时间维度上支持视频的延展) 	
万相 2.2	<ol style="list-style-type: none"> 1) 包含文生视频、图生视频、统一视频生成共四款模型 2) 文、图生视频模型为业界首个使用 MoE 架构的视频模型, 在同参数规模下, 节省约 50% 的算力消耗 3) 图生视频模型较万相 2.1 生成速度提升 12 倍, 抽卡成功率提升 123% 5) 显著提升复杂运动生成、人物交互、美学表达等 	<ol style="list-style-type: none"> 1) 电影美学控制系统 (可对光线类型、机位角度、高级运镜、视觉风格、特效镜头等实现控制, 光影、色彩、构图等能力媲美专业电影水平) 	

万相 2.5

- 1) 首次采用原生多模态架构，实现音画同步生成
- 2) 提升指令遵循能力，可理解运镜等复杂连续变化的指令

1) 音画同步生成功能



万相 2.6

- 1) 包含文生视频、图生视频、参考生视频共三款模型
- 2) 单次生成视频时长提升至 15 秒
- 3) 提升指令遵循能力、视觉质量、音频质量

- 1) 分镜头叙事功能（将提示词转换为多分镜头脚本，生成多镜头连贯叙事视频）
- 2) 角色扮演功能（参考视频的角色外观和音色实现角色扮演，生成单人、多人、人与物合拍视频）
- 3) 音频驱动生成功能（由文本和音频驱动视频生成，支持多镜头音频驱动叙事）



资料来源：阿里云官网，太平洋证券研究院

MiniMax 海螺：聚焦人物动作及表情表现力，推出细分场景视频模型。2024 年 8 月，MiniMax 发布海螺 1.0，强调稳定的运动生成，以及细致真实的人物情感表现。此后，海螺围绕人物动作及表情表现力持续迭代，海螺 2.0 及 2.3 持续提升复杂指令遵循和复杂物理表现能力，在人物肢体动作、面部表演、微表情变化等细节的呈现中优势显著。此外，海螺针对细分场景推出专用视频模型：1) 海螺 1.0-Live：专注于将静态 2D 图像转化为动态视频内容，聚焦二次元、插画风格的视频生成；2) 海螺 1.0-Director：专注于镜头运动的可控生成，提供 15 种可自由组合的单一运镜。

图表 17：MiniMax 海螺各代版本信息梳理

模型版本	发布时间	生成时长	分辨率	提示词类型	音频	定价
海螺 1.0	2024.08.31	6 秒	720p	文本、图片	无	1) 文/图生视频：0.5 元/秒 2) 主体参考生视频：0.75 元/秒
海螺 1.0-Live	2025.01.12	6 秒	720p	图片	无	0.5 元/秒
海螺 1.0-Director	2025.03.03	6 秒	720p	文本、图片	无	0.5 元/秒
海螺 2.0	2025.06.18	6/10 秒	512p、768p、1080p	文本、图片	有	1) 文/图生视频：768p：0.4 元/秒； 1080p：0.58 元/秒 2) 图生视频：512p：0.1 元/秒
海螺 2.3	2025.10.28	6/10 秒	768p、1080p	文本、图片	原生音频	1) 文/图生视频：768p：0.4 元/秒； 1080p：0.58 元/秒 2) Fast 图生视频：768p：0.225 元/秒； 1080p：0.385 元/秒

资料来源：MiniMax 官网，太平洋证券研究院

图表 18: MiniMax 海螺各代版本示例

模型版本	特点	新增功能	示例
海螺 1.0	1) 运动生成稳定性好, 人物情感表现细致真实 2) 可生成具备电影感的运镜效果、顶级的影视特效	1) 提示词优化功能 2) 单图主体参考功能 3) 创意模版功能 (预设创意模板, 将视频制作全流程自动化, 实现爆款内容快速复刻) 4) 上线海螺视频 App (拥有与网页版相同的功能, 提供作品发布分享的平台)	
海螺 1.0-Live	1) 专注于将静态 2D 图像转化为动态视频内容, 聚焦二次元、插画风格的视频生成	-	
海螺 1.0-Director	1) 专注于镜头运动的可控生成 2) 适用于多种视频创作场景, 为专业电影制作人、动画、广告等行业从业者、游戏开发者提供全新的创意施展空间	1) 镜头控制功能 (支持自然语言控制镜头运动, 提供 15 种可自由组合的单一运镜)	
海螺 2.0	1) 显著提升复杂指令遵循和物理表现 2) 具备价格优势, 每秒视频定价较 Runaway Gen4-Turbo、谷歌 Veo2、OpenAI Sora、可灵 2.1、豆包 Seedance 1.0 Pro 便宜	1) 首尾帧连接、尾帧生成功能 2) 音效生成功能 (通过集成音频生成模型, 生成和视频内容匹配的音效) 3) 上线海螺视频 Agent, 提供专业的视频创意 Agent 模版, 实现一键生成高质量创意短片	
海螺 2.3	1) 提升运动指令遵循和物理表现, 动作呈现更流畅自然 2) 更自然的真人面部表演、微表情变化 3) 升级动态表现力 (各种动态运镜下, 光线方向、色调等近乎实拍效果) 4) 擅长动漫、插画、水墨、游戏 CG 等特殊画风效果	1) 发布海螺 2.3 Fast 图生视频模型 (生成速度更快定价更低, 批量成本最高降低 50%) 2) 海螺视频 Agent 升级为支持全模态全能创作的 Media Agent (集成了分镜制作、剪辑、配音等视频制作全流程所需的工具, 可自动匹配并调用多模态模型, 一键生成高质量成片; 通过和 Agent 对话, 可介入其工作流并获得单独的片段素材, 或对内容进行编辑)	

资料来源: MiniMax 官网, 太平洋证券研究院

1.3 模型已支持 L3 短片级内容创作, 仍处技术快速迭代期

根据以上八个模型的版本迭代细节, 总结 AI 视频模型的演进趋势, 具体有以下五点:

画质与时长: 从低清短片向高清长视频演进, 海外重分辨率、国内重时长。海内外主流模型均将提升分辨率、延长时长、优化画质作为迭代方向之一, 并呈现海外优先突破分辨率, 国内侧重拓展单次生成时长的差异化路径。具体来看: 1) **原生分辨率最高为准 2K:** 海外模型领先, Sora2 原生分辨率最高, 为准 2K; Veo3、Ray3 可通过超分功能升至 4K; Ray3 支持生成 16bit-HDR, 已达电影行业对母版的精度要求。而国内模型最高支持原生 1080p。2) **单次生成时长最长为 25 秒:** 海外 Sora2 单次生成时长最长, 达 25 秒。国内万相 2.6 最先实现单次生成 15 秒, 可灵 3.0、Seedance 2.0

随后跟进，该时长长于除 Sora2 外的其他海外模型。同时，海内外模型均具备视频延长、拼接功能，可将时长延长至分钟级。

图表 19：海外视频模型优先突破分辨率，国内优先拓展时长

模型版本	发布时间	生成时长	分辨率
海外			
Sora2	2025.09.30	10/15/ 25 秒	720p、1080p、 准 2K (原生)
Ray3	2025.09.18	5/10 秒	最高 4K (超分)
Veo3.1	2025.10.15	4/6/8 秒	720p、1080p (仅限 8 秒)、 4K (超分)
Gen-4.5	2025.12	5/10 秒	720p
国内			
海螺 2.3	2025.10.28	6/10 秒	768p、1080p
万相 2.6	2025.12.16	2/5/10/ 15 秒	720p、1080p
可灵 3.0	2026.02.04	3-15 秒	720p、1080p
Seedance 2.0	2026.02.12	4-15 秒	480p、720p

资料来源：各公司官网，太平洋证券研究院

速度与成本：生成速度及成本持续优化，海外模型定价整体高于国内。海内外主流模型通过架构优化、发布轻量级模型等方式提升生成速度、降低生成成本，以适配 C 端创作者和 B 端创作团队的不同需求。具体来看：**1) 生成速度提升：**海内外模型均在迭代中实现不同程度的生成速度提升。如 Ray2 通过投入 10 倍于前代模型的算力，实现生成速度提升，万相 2.2 图生视频模型的生成速度提升 12 倍。**2) 生成成本间接降低：**一是发布轻量级模型，降低生成成本，如 Gen-4 的定价为 0.24 美元/秒，Gen-4 Turbo 的定价为 0.1 美元/秒。二是新版模型在提升生成质量的同时，维持定价不变，如 Veo、万相、海螺。**3) 海外模型定价整体高于国内：**对比各模型最新版本的定价，以文生、无声、无参考、不含视频输入、720p 视频定价为例，海外模型定价区间为 0.1-0.5 美元/秒，国内模型定价区间为 0.4-1 元/秒。

图表 20：主流视频模型生成速度及成本对比

模型版本		定价	生成速度及成本	轻量级模型
海外				
Sora	Sora Turbo	-	生成速度：根据创作者的测评数据，较预览版提高约 40%	-
	Sora2	0.1 美元/秒 (海外最低)	-	-
Veo	Veo3	0.4 美元/秒	-	-
	Veo3.1	0.4 美元/秒	-	Veo3.1 Fast
Gen	Gen-3 Alpha	0.2 美元/秒	-	Gen-3 Turbo (0.10 美元/秒)
	Gen-4	0.24 美元/秒	-	Gen-4 Turbo (0.10 美元/秒)
	Gen-4.5	0.5 美元/秒 (海外最高)	-	-
Ray	Ray2	\$0.14 美元/秒	算力资源是一代模型的 10 倍，实现生成速度提升	Ray Flash 2 (生成速度提升 3 倍、成本下降 67%)
	Ray3.14	-	1) 生成速度：提升 4 倍 2) 生成成本：下降 67%	-
国内				
可灵	可灵 3.0	0.6 元/秒	-	-
Seedance	Seedance 1.0 Pro	0.32 元/秒	-	Seedance 1.0 Pro Fast (生成速度最高提升约 3 倍，成本降低约 72%，0.09 元/秒)
	Seedance 2.0	1 元/秒 (国内最高)	-	-
万相	万相 2.2	0.4 元/秒	1) 生成速度：图生视频模型提升 12 倍 2) 生成成本：采用 MoE 架构，在同参数规模下，节省约 50% 的算力消耗	-
	万相 2.5	0.6 元/秒	-	-
	万相 2.6	0.6 元/秒	-	-
海螺	海螺 2.0	0.4 元/秒	-	-
	海螺 2.3	0.4 元/秒 (国内最低)	-	海螺 2.3 Fast 图生视频模型 (0.225 元/秒)

资料来源：各公司官网，太平洋证券研究院

多模态能力：输入端实现多模态混合，输出端实现音视频同出。在输入端，海内外主流模型通过创新多模态视觉语言交互范式、强化跨模态理解能力等方式，实现从仅支持文本或图像的单一模态输入，向支持文本+图片/音频/视频的多模态混合输入的升级，显著提升指令遵循、可控生成能力。**在输出端**，海内外模型均已实现音视频同步生成。具体来看，早在 2024 年 9 月，国内万相通过调用音频生成模型实现基础音效生成。随着多模态融合技术的成熟，2025 年 5 月，海外 Veo3 基于统一的多模态架构，率先实现原生音频与视频的同步生成，相较于早期的音频生成方式，显著提升音画同步精度。此后，国内主流模型自 2025 年 9 月起亦陆续完成音画同步生成的迭代。

图表 21：目前主流视频模型均已支持多模态输入、音画同步输出

模型版本	发布时间	提示词类型	音频
海外			
Veo3	2025. 05. 21	文本、图片	原生音频
Sora2	2025. 09. 30	文本、图片	原生音频
Ray3	2025. 09. 18	文本、图片、视频	无
Veo3. 1	2025. 10. 15	文本、图片、视频	原生音频
国内			
万相 1.0	2024. 09. 19	文本、图片	基础音效
海螺 2.0	2025. 06. 18	文本、图片	音效
万相 2.5	2025. 09. 24	文本、图片、音频	原生音频
海螺 2.3	2025. 10. 28	文本、图片	原生音频
可灵 3.0 Omni	2026. 02. 04	文本、图片、音频、视频	原生音频
Seedance 2.0	2025. 02. 12	文本、图片、音频、视频	原生音频

资料来源：各公司官网，太平洋证券研究院

可控性与适配性：从随机生成向精细化控制、专业化创作适配演进。海内外主流模型持续强化对人物、场景、镜头等创作要素的精细化控制能力，同时针对影视等专业场景进行优化，逐步适配专业创作流程。具体来看：**1) 全流程可控性提升：**在画面可控性上，海内外模型均实现角色、场景等的一致性提升，并围绕创作流程推出增强可控性的功能，如参考生成、镜头控制、运动笔刷等。在叙事可控性上，2025 年末起，国内万相 2.6、可灵 3.0、Seedance2.0 相继更新分镜能力，支持生成多镜头连贯叙事视频。**2) 适配细分和专业领域需求：**在内容生成上，海内外模型持续拓展美学风格覆盖广度，从基础风格向多元细分风格演进，满足多样化的创作需求，如 MiniMax 针对二次元、插画风格的视频生成推出海螺 1.0-Live。在专业场景适配上，Ray3 支持输出可供专业创作者编辑的 EXR 原始素材格式，直接适配后期制作。

图表 22：可灵 3.0 Omni 自定义分镜功能


资料来源：可灵 AI，太平洋证券研究院

图表 23：海螺 1.0-Live 模型


资料来源：MiniMax 官网，太平洋证券研究院

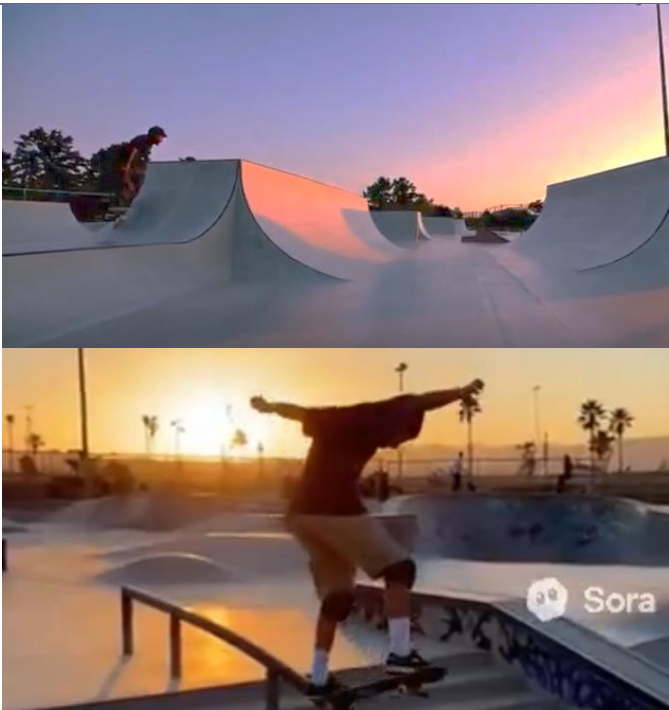
模拟能力：从生成简单画面向精准模拟物理规律、人物动作与表情演进。海内外主流模型从最初仅能实现基础的画面与动作生成，逐步向遵循物理规律、还原复杂人物动作与微表情升级，生成内容的保真度与合理性大幅提升。具体来看：**1) 遵循真实物理规律：**以 Sora 为例，根据 APIYI 数据，Sora2 的物理规律合规率为 88%，相较于前代模型 41%提升超一倍；人物行走、动态水流、烟雾扩散等具体情景的物理规律合规率相较于前代模型均实现提升。**2) 还原人物动作与表情：**以海螺为例，得益于模型在复杂物理表现与复杂指令遵循的进一步增强，最新的海螺 2.3 模型可呈现出舞蹈、跳跃等更复杂的人物肢体动作，以及更真实自然的真人面部表演、微表情变化等细节。

图表 24: Sora 和 Sora2 物理规律合规率对比

情景	物理规律合规率	
	Sora	Sora2
人物行走	65%	90%
动态水流	50%	85%
物体掉落	60%	92%
烟雾扩散	55%	88%
光影变化	70%	95%
织物飘动	58%	86%

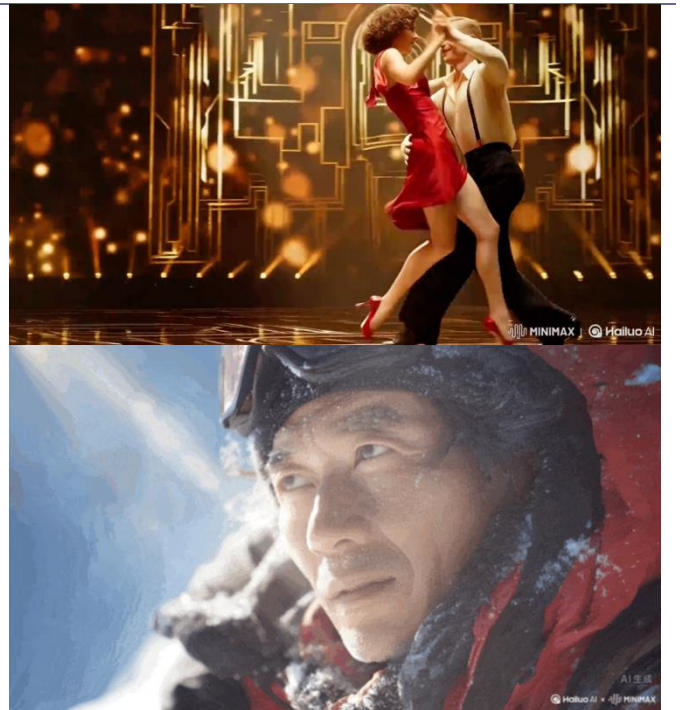
资料来源: APIYI, 太平洋证券研究院

图表 25: Sora 和 Sora2 生成滑板运动对比



资料来源: GenAI Intel, 太平洋证券研究院

图表 26: 海螺 2.3 生成复杂动作、人物表情



资料来源: MiniMax 官网, 太平洋证券研究院

基于“可控性、美学、物理”三大维度，评价模型内容生成能力。AIGC 内容创作分级标准可依据可控性、美学风格、物理模拟三大维度，将 AIGC 内容创作划分为五个级别，用于评价视频模型的能力等级。级别越高，代表模型支持生成的视频质量越高、时长越长。L1 级至 L5 级覆盖从基础素材到专业影视内容，即 L1 对应 5-10 秒的单镜头和分镜素材，L2 对应数十秒的短视频片段，L3 对应数分钟、带叙事结构的短片，L4 对应 20 分钟以上的中长完整剧情视频，L5 对应消费级影视内容。

图表 27: AIGC 内容创作分级标准

	可控性	美学	物理		
	提示词和图形引导控制 少量可调参数(提示词相关性等)	支持首尾帧衔接，带来基础编排能力 平移、垂直、推进等基础运镜	角色形象、口型、动作基本可控 平移、垂直、推进等基础运镜	复杂角色行为与场景可控(精确行为控制) 复合运镜、多人同屏互动，多关键帧衔接	全流程可控，高精度动作和表情控制、多角色同屏调度，物理模拟和3D镜头控制精细化
	运镜控制(基础) 首尾帧衔接 关键	口型匹配 关键 基础表情 运动笔刷 主体参考 角色一致性 关键 运镜控制(基础) 首尾帧衔接	局部控制与重编辑 口型匹配 复杂表情 运动笔刷 主体行为控制 关键 多主体参考 同屏角色一致 关键 运镜控制(影视级) 多关键帧衔接 关键	虚拟场景控制系统 关键 场景生成与一致性、局部精细控制与编辑、角色场景联动控制等 虚拟表演系统 关键 多IP锁定、口型生成、表情生成、精确角色行为控制等 三维虚拟运镜控制 关键 多镜头生成与自动剪辑 关键	
	基础风格化质量，风格类型少	更多主流风格支持：二次元、3D动画、漫画等	更多品类细分和风格滤镜支持，风格稳定性和区分度提升	进入专业分镜和完整叙事场景，对整体氛围(灯光、调色、构图)更高要求	实现影视工业级别的调色、光影、特效统一；支持复杂而多变的艺术风格
	美学风格化L1	美学风格化L2	美学风格化L3	美学风格化L4	美学风格化L5
	几乎无物理模拟，微动/缓动视频	基础动作：行走、举手、跑步等画面连续不破碎	基础物理模拟：重力、碰撞等常规动作和连续复合动作	更复杂的物理模拟，流体、布料、复杂碰撞预测和复杂人体运动	世界模拟器级别的高精度、多维度物理模拟，角色与环境互动逼真，场景稳定
	文生视频L1 图生视频L1	文生视频L2 图生视频L2	文生视频L3 图生视频L3	文生视频L4 图生视频L4	文生视频L5 图生视频L5
	L1 单镜头 5-10秒单镜头和分镜素材生产	L2 片段级 数十秒短视频剪辑片段	L3 短片级 可生产数分钟带叙事结构的短片	L4 长片级 中长(20分钟+)完整剧情视频	L5 影视级 应用到消费级影视制作 workflow

资料来源：阿里云，短剧自习室，太平洋证券研究院

AI 视频模型目前已支持 L3 短片级内容创作，并部分满足 L4 长片级标准。基于 AIGC 内容分级标准，并结合上文总结的目前 AI 视频模型的五大演进趋势，我们认为，目前 AI 视频模型已支持 L3 短片级内容创作。即：1) 在可控性方面，达到角色形象、口型、动作基本可控，且拥有平移、垂直、推进等基础运镜；2) 在美学风格方面，支持写实、电影、动漫等多种风格；3) 在物理模拟方面，可对基础物理规律进行模拟，且可呈现常规动作和连续复合动作。此外，模型已部分满足 L4 长片级内容创作的标准，如实现多主体参考、影视级运镜控制等。

2. 影视场景：细分领域渗透率提升，降本增效驱动增长

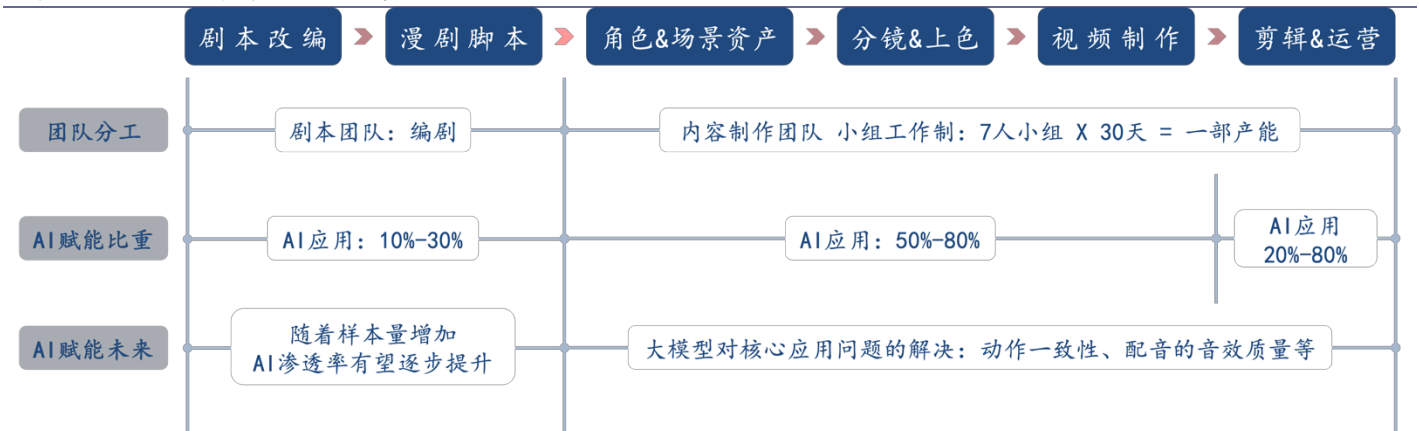
视频模型 L3 短片级内容制作能力融入影视 workflow，推动行业迈入 AI 普及期。目前 AI 视频模型支持 L3 短片级内容创作，部分能力达到 L4 长片级内容要求，一定程度上已满足影视行业制作需求。同时，视频市场全球空间广阔，AI 渗透率仍处低位，技术持续落地将推动 AI 渗透率快速提升。但目前模型仍较难直接适配影视领域专业 workflow，需与制片流程深度融合。因此，以 AI 视频模型为基础，集成角色及场景设计、智能剪辑等专业功能的 AI 视频工具同步加速迭代。

2.1 影视细分领域：AI 漫剧率先落地，AI 仿真人剧快速起量

目前，由 AI 辅助制作或完全由 AI 生成的 AI 原生影视作品，已在漫剧、短剧、电影和电视剧等领域崭露头角。

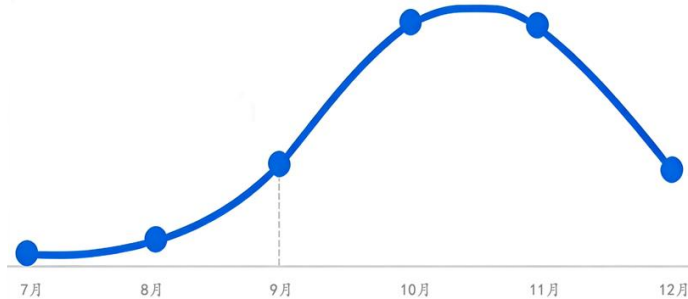
1) 漫剧：AI 赋能全流程制作，漫剧供给迎来爆发式增长。2025 年，AI 生成技术已实现漫剧制作全流程赋能，形成 AI 漫剧的成熟 workflow：1) 剧本创作环节：AI 应用比例为 10%-30%，主要用于 IP 剧情理解和大纲拆解。未来随着样本量增加，AI 渗透率还将逐步提升。2) 内容制作环节：AI 应用比例为 50%-80%，主要用于角色、场景、分镜等内容的生成，以及后期剪辑。未来随着模型一致性、叙事可控性等能力提升，生成质量有望持续突破。得益于 AI 赋能全流程制作，2025 年漫剧行业迎来供给爆发式增长，AI 漫剧是核心驱动力。根据巨量引擎，2025 年 9 月后抖音漫剧供给数量持续提升，截至 11 月底，突破 10 万部，增速达 59%。同时，2025Q4 AI 漫剧在漫剧榜单中的占比已超 70%。

图表 28：AI 漫剧成熟实践 workflow



资料来源：巨量引擎，太平洋证券研究院

图表 29：2025 年 7-12 月漫剧供给数量



资料来源：巨量引擎，太平洋证券研究院

AI 赋能下，动态漫为发展最成熟的漫剧类型，现象级作品实现低成本撬动高收入。根据画风和动态精细度的不同，漫剧主要包括表情包漫剧和动态漫两种类型。根据 DataEye-ADX，2026 年 1 月在播漫剧中，2D、3D 漫剧的占比达 71%，远高于表情包漫剧的 15%，是目前漫剧行业最成熟的内容形式。动态漫具有代表性的现象级作品为《斩仙台下，我震惊了诸神！》，该剧是抖音播放量首部突破 10 亿的漫剧作品。其实现了从情节爽点提取到角色建模、动画生成以及后期剪辑的全流程 AI 赋能，由 10 人团队在 10 天制作周期内，以 15 万元的成本完成制作。根据巨量引擎数据，截至 2026 年 1 月，该剧在红果短剧及抖音平台上的累计分账收入达 140 万元。

图表 30：主流漫剧分类

类型	简称	核心概念	特点	示例
表情包漫剧	沙雕漫	以互联网中广泛流通的各类表情包、网络梗图等作为核心视觉素材，通过编排、动态化与配音，组合成有连续剧情的内容。	依托表情包自带的情绪张力和大众认知基础，快速构建角色、传递情绪，是一种成本低廉、风格诙谐、易于传播的轻量内容形态。	
动态漫画	动态漫	平面漫画与动态元素相结合的动画表现形式，一般是将传统静态漫画经过技术处理之后，转变为一种具有动态效果的动画作品。	在漫画图片的基础上进行动态处理，令镜头推拉摇移，漫画人物或事物能做出向前走、电梯门拉开等简单动作，辅以对话和背景音乐等。	

资料来源：短剧自习室，太平洋证券研究院

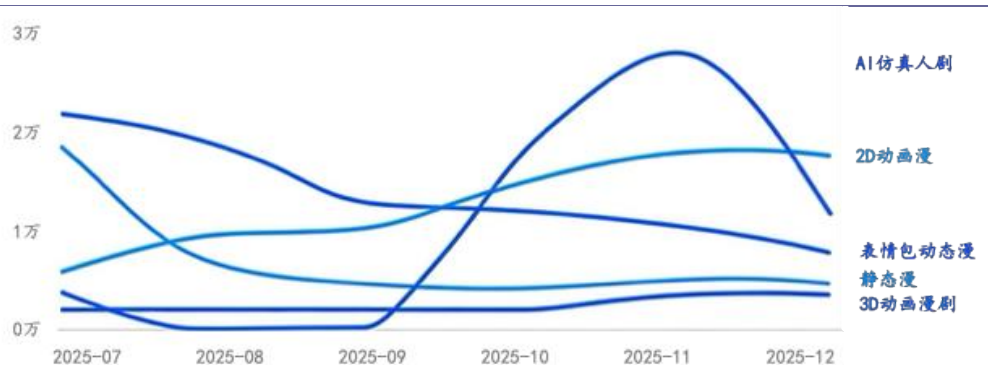
图表 31: 《斩仙台下，我震惊了诸神!》



资料来源: 巨量引擎, 太平洋证券研究院

2) 真人短剧: 由“AI+实拍”发展至全 AI 制作, AI 仿真人剧快速起量。随着 AI 技术的逐渐成熟, 其在真人短剧制作中的应用已从早期相对保守的赋能具体环节, 逐步发展至全流程生成。相应的, AI 短剧从早期的“AI+实拍”形式, 发展至 AI 仿真人剧。根据巨量引擎数据, 播放量方面, AI 仿真人剧的部均播放量增速最快, 且在 2025 年 11 月达到峰值, 高于 2D、3D 动画漫、表情包动态漫等。供给方面, 产量有所增加, 12 月上新 50 部。投流方面, 自 2025 年 11 月开始起量, 12 月月月在投 900 部, 目前已有末世生存、悬疑诡秘等题材收获亮眼的播放数据, 带动投流千万消耗增量。

图表 32: 2025. 07-2025. 12 抖音各类型短剧部均播放量



资料来源: 巨量引擎, 太平洋证券研究院

AI 仿真人剧质量持续提升, 头部作品播放量快速破亿。早期 AI 真人短剧以“AI+实拍”的方式完成制作, 以古风奇幻题材短剧《浮梦吟》为例, AI 制作内容在全剧中

的占比约为 15%。随着模型能力持续提升，完全由 AI 生成的 AI 仿真人剧崭露头角。以可梦 AI 相关作品为例，2025 年 8 月和 10 月、2026 年 1 月可梦 AI 分别上线了宫廷短剧《奶团太后宫心计》、年代歌舞短剧《我靠唱歌打脸全团》和神话仙侠短剧《斩仙台 AI 真人版》，三部作品在人物表情控制、画面特效丰富度等方面均实现显著提升。其中《斩仙台 AI 真人版》上线仅 6 天抖音播放量破亿，是目前播放量破亿最快的 AI 仿真人剧。但其表现与真人实拍短剧仍存在一定差距，如《家里家外》上线 3 天红果短剧播放量破 10 亿。此外，《我靠唱歌打脸全团》和《斩仙台 AI 真人版》均跑通了商业化，ROI 突破 1.1。

图表 33: AI 仿真人剧代表作品

奶团太后宫心计	我靠唱歌打脸全团	斩仙台 AI 真人版
		
<ul style="list-style-type: none"> 上线: 2025.08.08 团队: 10人/15天 成本: 仅为传统短剧的1/4 热度: 女频类漫剧第一 播放: 2.2亿+ ROI: - 	<ul style="list-style-type: none"> 上线: 2025.10.03 团队: 12人/10天 成本: - 热度: 连续霸榜7天 播放: 1.17亿+ ROI: 1.1+ 	<ul style="list-style-type: none"> 上线: 2026.01.13 团队: 12人/30天 成本: 10万 热度: AI短剧/红果动态榜第一 播放: 1亿+ ROI: 1.1+

资料来源: 可梦 AI, 太平洋证券研究院

3) 电影、电视剧: 受限于模型能力, AI 应用目前仍以辅助工具为主。目前 AI 在电影、电视剧的应用仍面临生成画质较难满足行业要求, 人物微表情、肢体语言的情感传递能力有限等问题。同时, 观众对电影、电视剧的质量要求相较短剧更高, 因此目前 AI 在电影、电视剧, 尤其是真人内容领域的应用仍以辅助工具为主。包括: AI 换脸、辅助大成本场景及特效生成等。如: 由博纳影业出品的奇幻动作电影《传说》利用 AI 技术重现 27 岁的成龙形象, 奈飞出品的阿根廷科幻剧集《El Eternauta》利用 AI 技术生成布宜诺斯艾利斯建筑倒塌的灾难画面等。

图表 34: AI+短剧和 AI+电影、电视剧对比

对比形态	AI+短剧	AI+电影、电视剧
创作门槛	适中，具备高于短视频叙事与制作的专业门槛	未达到可商用的技术落地水平
商业模式	2025 年年中跑通商业闭环，较易变现	未有实际累计票房或收视率的落地案例，传统 B 端商业模式与 AI 快速迭代逻辑不完全匹配
供给表现	成本较低，验证周期快	成本高，验证周期漫长
消费体验	用户容忍度高，对内容质量有一定要求	用户容忍度低，对内容质量要求最高
产品特征	典型的数据驱动产品，用户数据反馈高频且密集，可快速验证与优化	内容驱动产品，用户数据反馈稀疏且缓慢，难以快速验证与优化
画质水平	AI 生成的画质逐步提升，已能满足短剧的视觉化呈现	部分 AI 模型已能支持 4K 内容的生成，但生成影视剧等长片内容的效果暂未成熟

资料来源：短剧自习室，太平洋证券研究院

图表 35: 《传说》27 岁成龙形象



资料来源：《传说》预告片，太平洋证券研究院

图表 36: 《Eternauta》AI 画面



资料来源：《Eternauta》，太平洋证券研究院

AI 动画电影率先落地，真人电影仍处起步阶段。相较于真人电影对写实性的极高要求，动画电影的风格化表达为 AI 生成提供了容错空间，AI 生成更易实现其对角色面部表情、动作设计等内容的创作。因此，目前 AI 动画电影相较于 AI 真人电影发展更成熟。**动画电影方面**，中国首部全流程 AIGC 动画电影《团圆令》于 2 月 28 日正式上映，该影片采用 AI 全流程动画生产平台、历时 5 个月完成制作；中国首部全 AI 院线动画电影《山海经·玄武》预计于 2026 年暑期档上映。**真人电影方面**，2024 年 3 月，全球首部 AI 长篇电影《Our T2 Remake》在洛杉矶上映；2025 年 4 月，全球首部院线公映的 AIGC 电影《海上女王郑一嫂》在新加坡上映。从影片画面来看，目前 AI

真人电影仍存在人物表情缺乏情感传递、缺乏叙事逻辑等不足，与真人实拍电影存在较大差距。

图表 37：中国首部全流程 AIGC 动画电影《团圆令》



资料来源：《团圆令》，太平洋证券研究院

图表 38：中国首部全 AI 院线动画电影《山海经·玄武》



资料来源：《山海经·玄武》，太平洋证券研究院

图表 39：全球首部 AI 长篇电影《Our T2 Remake》



资料来源：《Our T2 Remake》，太平洋证券研究院

图表 40：全球首部院线 AIGC 电影《海上女王郑一嫂》



资料来源：《海上女王郑一嫂》，太平洋证券研究院

2.2 AI 赋能降本增效，渗透率仍有较大提升空间

基于以上对漫剧、短剧、电影和电视剧等领域的梳理分析，我们认为，AI 在视频制作 workflows 中的深度融合，正推动视频行业实现降本增效，驱动行业增长。

1) 降本：AI 驱动全流程降本，素材拍摄环节最受益

AI 驱动视频制作全流程降本。目前 AI 已对视频前、中、后期制作的各个环节实现降本。1) **前期制作：**AI 可替代编剧、分镜师，自动生成大纲、剧本、分镜等内容，显著降低人力成本。根据短剧自习室，AI 分镜日均可生成 50 集标准化分镜，效率远

高于人工分镜师月均 8-10 集。2) 中期制作：AI 生成替代实景拍摄，其生成的人物形象、场景等均为数字资产，具备稳定可复用性。这将使得单一场景、人物角色可在同系列作品中持续复用，且可有效规避真人演员违法失德导致作品无法正常播出等风险，降低拍摄设备及场地租赁、演员片酬等支出、规避因演员风险带来的沉没成本等。3) 后期制作：AI 智能剪辑等技术可大幅减少剪辑师的人数和工时。此外，模型音视频同出的能力，使得音效环节无需单独配音，减少后期制作成本。

图表 41：AI 赋能视频制作全流程降本

	大纲脚本	分镜规划	素材拍摄	素材选取	剪辑	渲染、音效
主要内容	<ul style="list-style-type: none"> 故事创意、主题风格构思，明确故事大纲，如主要情节和角色发展 将大纲扩展成剧本，包括对白、场景描述和情节细节 	<ul style="list-style-type: none"> 剧本视觉化，确定镜头类型及运动、角色位置和动作 编制镜头列表，包括每个镜头的编号、描述、拍摄角度、镜头运动等 	<ul style="list-style-type: none"> 根据分镜规划制定拍摄计划，完成器材准备和技术测试、准备演员、服化道及布景设置 	<ul style="list-style-type: none"> 根据情节和主题需要对素材整理、分类、剔除无效素材 对于无须拍摄类内容，需要从素材库进行素材选取 	<ul style="list-style-type: none"> 初步剪辑、逻辑剪辑，拼接出粗略的视频版本 精剪微调，对影片进行精细剪辑，调整镜头长度、顺序、节奏等 	<ul style="list-style-type: none"> 添加视觉特效，确保效果自然、完成颜色校正、调色与整体风格一致 同步处理音频素材，确保对白清晰、搭配音效，增强视频的听觉效果
AI 降本	<ul style="list-style-type: none"> AI 替代编剧、分镜师，自动生成大纲、剧本、分镜等内容 		<ul style="list-style-type: none"> AI 生成替代实景拍摄，生成人物角色、场景画面、特效等内容 		<ul style="list-style-type: none"> AI 智能剪辑等可根据节奏、音乐卡点剪辑 模型实现音视频同出，使得音效环节无需单独配音 	

资料来源：量子位智库，太平洋证券研究院

素材拍摄环节因成本占比高、AI 替代潜力大，降本效应最为显著。以短剧行业为例，根据峯野掌镜人对爆款短剧的统计数据，制作拍摄环节成本在总成本中的占比最高，为 40%-50%。同时，根据短剧自习室，目前竖屏短剧头部演员的日薪已达 3-8 万元，相较 2024 年的 1-3 万元实现翻倍上涨，进一步推高短剧制作成本压力。而 AI 对素材拍摄环节具备较大的替代潜力，其可实现人物角色、场景等核心制作素材的全面数字化。在制作拍摄成本占比高、短剧演员片酬大幅上涨的行业背景下，AI 赋能短剧降本效应最显著的环节在于素材拍摄。

图表 42：短剧成本构成及占比

制作阶段	成本占比	内容
前期开发	15%-20%	剧本定制、场景勘景
制作拍摄	40%-50%	演员片酬、设备租赁
后期制作	20%-25%	-
宣发投流	15%-20%	-

资料来源：峯野掌镜人，太平洋证券研究院

得益于成熟的 AI workflow，AI 漫剧成本较传统漫剧已实现大幅下降。相较于传统动漫短剧，AI 在漫剧制作中的应用可显著降低内容成本与人力成本。

1) 内容成本：传统动漫短剧需经过手工绘图、逐帧制作、后期合成等复杂流程，内容成本较高，而 AI 漫剧依托 AI 生成技术直接输出画面内容，大幅降低手绘、动画制作、渲染等内容成本。根据巨量引擎，AI 赋能可使漫剧的生产成本降低超 70%。

2) 人力成本：AI 在分镜生成、素材生成、剪辑等环节的成熟应用，使得大量原本依赖人工的工作被 AI 替代，带来团队规模大幅精简。相较传统动漫短剧需超百人的制作团队，AI 漫剧仅需数人至十余人即可完成制作，人力成本大幅下降。目前，AI 漫剧的制作成本为 1000-2500 元/分钟，制作难度、质量相对更低的沙雕漫的制作成本为 200 元/分钟，相较于传统动漫短剧数万至十数万元/分钟的制作成本，实现大幅下降。

图表 43：传统动漫短剧和 AI 漫剧制作成本、团队人数对比

	传统动漫短剧	AI 漫剧
制作成本	3-5万元/分钟 质量对标国漫：十数万/分钟	漫剧：1000-2500/分钟 沙雕漫：200+/分钟
团队人数	超百人	数人 - 十余人

资料来源：巨量引擎，短剧自习室，太平洋证券研究院

2) 增效：AI 显著压缩内容制作周期，丰富内容题材及形式供给

效率方面：AI 从生产关系、制作流程与可控性方面，优化视频制作流程。人与 AI 协作提高沟通效率：传统影视制作依赖导演、摄影等多岗位紧密配合，多环节的信息传递需耗费极大的沟通成本。AI 赋能后，导演可通过指令直接生成视频，提高沟通效率。

并行可回溯流程提升制作效率：传统影视制作采用顺序 workflow，剧本创作、角色及场景设定、拍摄等环环相扣，且前序流程更改牵动后续流程更改会产生较大成本。AI 赋能后，各环节可同时进行，且可反复回溯调整。

指令化调控提高制作可控性：传统影视制作中的画面、光影、运镜等高度依赖导演、摄影的现场协作，在多环节、多人员的创作环境中，可控性较弱。AI 赋能下，可通过文本指令、参数设定等实现精准调控，有效降低人为偏差。

图表 44：传统及 AI 影视制作流程对比



传统流程

- | | | |
|--|---|---|
| <ul style="list-style-type: none"> • 生产关系：人与人 1) 制作团队人数多 2) 人员沟通效率低 | <ul style="list-style-type: none"> • 制作流程：顺序工作流 1) 各环节环环相扣 2) 各环节费用在数万-数十万，返工成本大 | <ul style="list-style-type: none"> • 制作可控性：易产生偏差 1) 高度依赖人工经验与团队协作 2) 在多环节、多人员协作中，易出现理解偏差 |
|--|---|---|



AI 赋能流程

- | | | |
|---|--|--|
| <ul style="list-style-type: none"> • 生产关系：人与AI 1) 通过文本指令直接转化为画面 2) 提高沟通效率、减少沟通损耗 | <ul style="list-style-type: none"> • 制作流程：并行可回溯 1) 各环节可同时进行 2) 制作中可随时回溯调整 | <ul style="list-style-type: none"> • 制作可控性：精准可控 1) 通过文本指令、参数设定等精准生成 2) 降低人为偏差 |
|---|--|--|

资料来源：短剧自习室，太平洋证券研究院

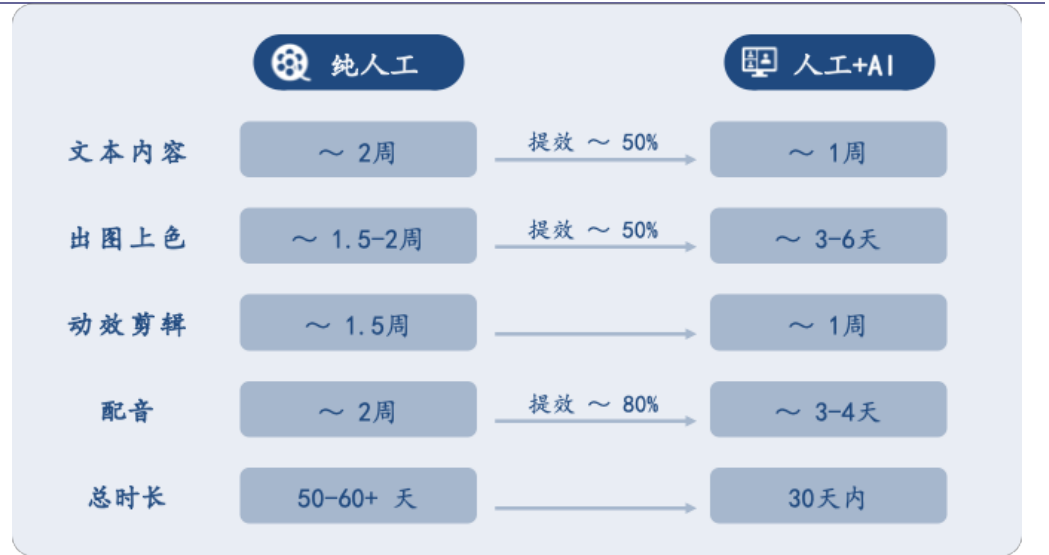
得益于流程优化推动效率提升，漫剧制作周期已压缩至周、月级别。以漫剧行业为例，传统动漫短剧的制作周期以年为单位，AI 对制作流程的优化推动漫剧制作周期已压缩至以周、月为单位。具体来看，漫剧中的文本内容、出图上色等制作环节在 AI 赋能下实现 50%-80%的提效，这使得 AI 赋能下的漫剧制作能在 30 天内完成，显著低于纯人工制作的 50-60 天。此外，AI 应用带来的降本增效亦在电影行业体现。根据公开数据，中国首部全流程 AIGC 动画电影《团圆令》的制作成本降低约 30%，制作周期约 5-6 个月，约为传统动画电影 2-3 年制作周期的五分之一。

图表 45：传统动漫短剧和 AI 漫剧制作效率、周期对比

	传统动漫短剧	AI 漫剧
制作效率	人工分镜师：8-10集/月	AI分镜：50集标准化分镜/天
制作周期	以年为单位 单条视频约15天	以周/月为单位

资料来源：短剧自习室，太平洋证券研究院

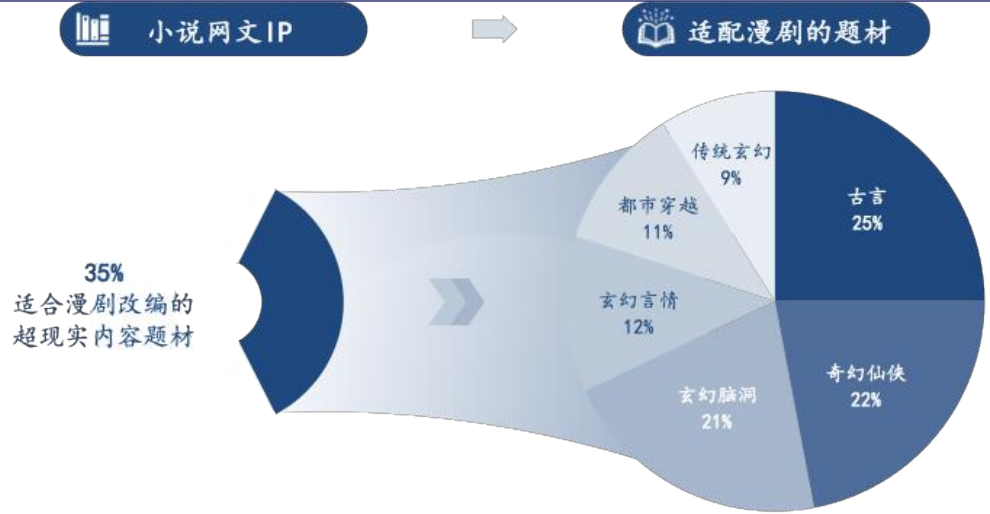
图表 46：纯人工和人工+AI 制作漫剧效能对比



资料来源：巨量引擎，太平洋证券研究院

供给方面：题材上，AI 赋能超现实题材漫剧制作，补充短剧题材空白。根据巨量引擎，网文 IP 中超现实内容题材，如奇幻仙侠、玄幻脑洞等的占比约为 35%。这些内容题材因受制于制作成本高、周期长等问题，改编成短剧的难度较大，目前尚未被充分开发。相较于对真实性要求更高的真人短剧，目前视频模型与动画、漫画风格的漫剧制作适配度更高。同时，得益于 AI，漫剧制作成本及周期已实现大幅优化。因此，我们认为，超现实题材内容有望在 AI 赋能下改编为漫剧，丰富短剧行业内容供给。根据短剧自习室及 CSM 统计数据，2025 年漫剧题材播放量前三为架空玄幻、逆袭、奇幻，2025 年 1-9 月热播热议微短剧题材前三为情感、古装、都市。

图表 47：网文 IP 内容题材结构



资料来源：巨量引擎，太平洋证券研究院

供给形式上，AI 通过 IP 复用与品类拓展，推动存量短剧实现内容形式创新。以 AI 转绘漫剧为例，其通过 AI 视频风格迁移技术，将存量真人短剧转绘为漫剧，满足年轻二次元用户观看需求，并有效激活存量真人短剧，实现爆款 IP 在真人短剧和漫剧的复用，进一步释放 IP 价值。根据短剧自习室，《办公室生存指南》经 AI 转绘为 Q 版职场番后，18-25 岁观众占比从 12% 提升至 58%，并带动衍生周边销售增长。根据巨量引擎，近一年商业化表现较好的存量真人短剧改编为漫剧的供给空间在 500-800+ 部。

图表 48：多样化内容形式引入



资料来源：巨量引擎，太平洋证券研究院

图表 49：AI 转绘漫剧《心动的信号》

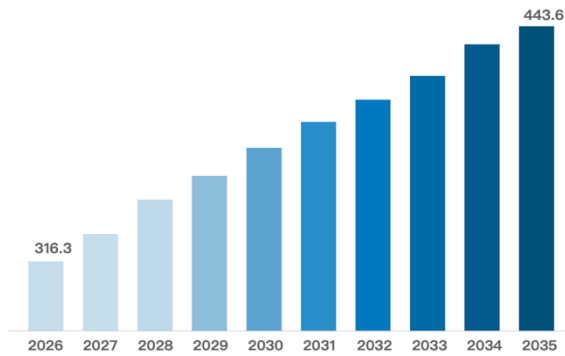


资料来源：红果短剧，太平洋证券研究院

3) 市场空间：全球潜在市场近 3200 亿美元，AI 视频渗透率仍有较大提升空间

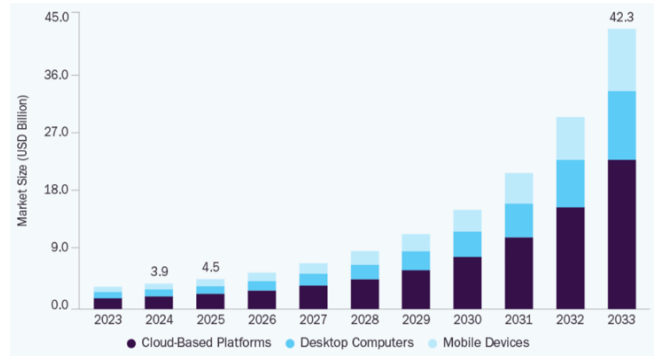
2025 年全球 AI 视频市场约 45 亿美元，渗透率仍处个位数。视频制作市场是 AI 视频生成可触达的潜在市场。从潜在空间来看，根据 Business Research Insights 预测，2026 年全球电影及视频制作市场规模约 3163 亿美元，覆盖广告、动画、电影、电视节目等内容。从渗透率来看，根据 Grand View Research 预测，2025 年全球 AI 视频市场规模约 45 亿美元，对应渗透率约 1.4%，处于起步阶段。到 2033 年，市场规模有望增长至 423 亿美元，CAGR 约 32%，对应渗透率约 10%，仍处较低水平。我们认为，更长期去看，假设 AI 视频数量占比为 70%，同时各环节 AI 视频工具应用比例为 70%，则远期 AI 视频渗透率有望达 50%。

图表 50：2026 全球电影及视频制作市场 3163 亿美元



资料来源：Business Research Insights, 太平洋证券研究院

图表 51：2023-33 全球 AI 视频市场规模(十亿美元)



资料来源：Grand View Research, 太平洋证券研究院

AI 漫剧率先贡献增量，AI 真人短剧渗透率快速提升，电影、电视剧有望打开更大的市场空间。目前，AI 漫剧已成为 AI 内容产业发展最成熟的细分领域，贡献 AI 视频市场规模增量。我们认为，随着 AI 视频模型持续迭代，AI 内容创作将从动漫向真人内容扩圈：**1) 受众广度：**AI 漫剧受众为泛二次元用户，具有明确圈层属性；而 AI 真人短剧受众覆盖近乎全民，具有更广泛的情感共鸣基础与消费习惯。**2) 市场规模：**2026 年中国漫剧市场规模约 220 亿元；而 2025 年中国真人短剧市场规模已近 750 亿元。随着 AI 仿真人剧的快速起量，AI 在真人短剧制作环节的渗透率有望迅速提升。**展望未来，**AI 电影、电视剧有望打开更大的市场空间。

图表 52: AI 漫剧和 AI 真人短剧对比

对比维度	AI 漫剧	AI 真人短剧
受众广度	核心与潜在受众基于超 5.26 亿的泛二次元用户，虽规模庞大，但仍有明确的圈层属性。	受众覆盖近乎全民，具有更广泛的情感共鸣基础与消费习惯。
国内市场规模	远小于 220 亿元 巨量引擎预估，2026 年中国漫剧市场 220 亿元，AI 仿真人剧的增长和占有率预计占比不小，所以 AI 漫剧在国内远小于 220 亿元。	大于 700 亿元 DataEye 的《2025 微短剧年度报告》显示，2025 年中国真人微短剧市场规模接近 750 亿元。

资料来源：短剧自习室，太平洋证券研究院

3. AI 视频标的梳理：视频工具为核心，IP 环节受益最大

AI 视频工具为产业链核心，视频模型、IP 及影视、第三方工具公司参与布局。AI 视频工具是模型能力转化为实际生产力的载体，以视频模型为基础，集成角色及场景设计、智能剪辑等专业内容制作功能。因既要依赖底层模型技术支撑，又需专业视频内容制作能力，所以目前 AI 视频工具领域已形成以互联网公司为代表的视频模型厂商，以及应用端 IP 及影视公司协同发展的格局。此外，第三方公司凭借 AI 工具开发技术与垂类内容制作经验，亦相继推出 AI 视频工具。我们认为，未来拥有技术领先优势的 AI 视频工具公司更易依托自身创作能力与平台生态，产出优质视频内容。同时，IP 公司因具备海量可视频化的内容库资源，有望充分受益于 AI 视频工具的成熟。

图表 53: AI 视频工具赛道三类公司

	视频模型公司		IP 及影视公司		第三方工具公司		
公司	快手	字节跳动	博纳影业	中文在线	智灵新境	可梦智能	灵境万维
应用	可灵 AI	即梦 AI	博乐	次元神笔	新境 AI	可梦 AI	灵境 AI
优势	技术壁垒高、迭代快 商业化场景丰富，覆盖电商、营销等		IP 转化与影视内容落地能力强 可将一线制作需求融入 AI 工具的迭代		兼具 AI 工具开发的技术实力 与垂类内容制作的实操经验		

资料来源：太平洋证券研究院整理

3.1 视频模型公司：占据技术和商业化场景优势，处主导地位

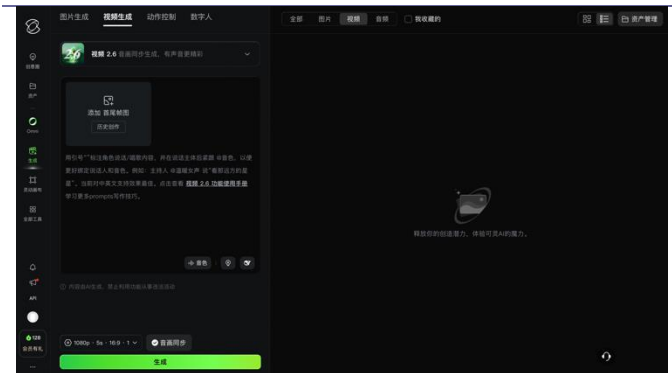
视频模型公司依托底层模型技术、算力、生态流量等优势，推出可灵 AI、即梦 AI 等通用 AI 视频工具。相较于 IP 及影视、第三方工具公司，其凭借技术壁垒高、迭代快、商业化场景丰富等优势在 AI 视频工具领域占据主导地位。

1) 快手：可灵 AI 构建全场景 AI 视频生产力，“工具+内容”共促 AI 视频发展

AI 视频工具: 1) 快手推出全场景 AI 视频生产力引擎可灵 AI, 该平台基于自研可灵视频、可灵图片模型, 面向影视制作、电商营销、数字人直播、社交内容等产业应用与个人创作场景, 提供文生视频、图生视频、AI 换装等专业功能。得益于可灵 o1 发布, 及可灵 2.6 动作控制功能在社交媒体引起广泛讨论, 2025 年 12 月可灵 AI 单月收入超 2000 万美元, 2026 年 1 月月活跃用户突破 1200 万。2) 快手推出一站式漫剧创作工具造梦专家, 实现将一段文字高效转化为一个拥有定制化角色、场景和动态画面的完整视频剧集。

AI 内容: 可灵 AI 先后举办了导演共创计划、NEXTGEN 全球新影像创作大赛等 AIGC 内容创作活动, 推动 AI 视频从技术演示走向落地应用。活动中诞生了许多的优秀作品, 如《新世界加载中》: 全球首部 AI 单元剧集, 在北京国际电影节 AIGC 单元获得最佳技术奖。《Kling 汉堡》: 可灵 AI 视频大赛广告天才主题冠军等。此外, 快手联合可灵 AI, 通过现金激励、流量扶持、联合出品及运营等方式, 扶持 AI 漫剧与短剧创作者, 推动内容商业化落地。

图表 54: 可灵 AI 工作台



资料来源: 可灵 AI, 太平洋证券研究院

图表 55: 全球首部 AI 单元剧集《新世界加载中》



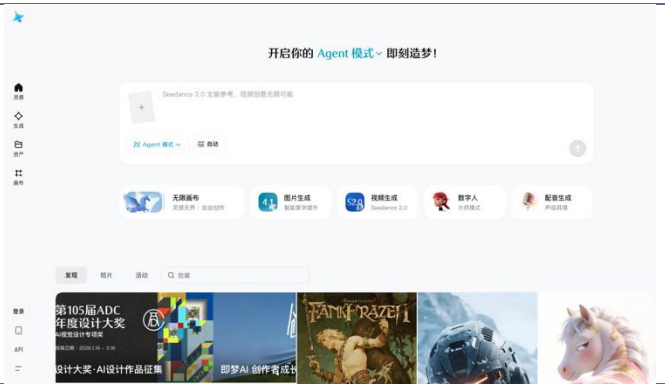
资料来源: 可灵 AI, 太平洋证券研究院

2) 字节: 即梦 AI 提供全链路一站式 AI 创作, 各类政策支持 AI 视频商业化落地

AI 视频工具: 字节跳动推出一站式 AI 创意创作平台即梦 AI, 该平台依托自研 Seedance、Seedream 等模型, 提供文生图、文生视频、数字人制作等功能, 支持从灵感搜索、到创意设计, 再到生成编辑成片的全链路创作, 满足个人与商业场景的视觉内容创作需求。根据量子位智库, 2026 年 2 月即梦 AI 网页版月访问总量超 1600 万, 环比增长 73%。根据 Inshtrackr, 海外版 Dreamina AI 下载量在 2025 年 12 月中下旬迎来显著增长。截至 12 月 26 日海外 12 国累计下载量约 29.5 万, 主要系 Seedance 1.5 Pro 等发布, 以及 AI 内容在社交媒体广泛传播。

AI 内容：即梦 AI 先后举办 AIGC 短剧招募计划、青年导演合作计划等 AIGC 内容创作活动，加速短剧、艺术短片等的工业化生产。活动中诞生了许多的优秀作品，如《老妈的心愿》：家庭温情 AI 短片，抖音上线 24 小时收获点赞超 3 万。《金鱼》：获得 2025 年上海国际电影节 AIGC 竞赛单元最佳视效奖等。此外，字节跳动依托即梦 AI 与抖音、红果生态，通过 IP 开放、技术成本补贴、分级保底等方式，全面扶持 AI 漫剧与短剧内容创作与商业化落地。

图表 56：即梦 AI 工作台



资料来源：即梦 AI，太平洋证券研究院

图表 57：《金鱼》



资料来源：即梦 AI，太平洋证券研究院

3.2 IP 及影视公司：将一线需求融入工具迭代，赋能 AI 影视工业化

IP 及影视公司依托影视内容制作经验，自研适配于影视工业化生产的 AI 视频工具。相较于视频模型、第三方工具公司，其 IP 转化与影视内容制作经验深厚，可将一线制作需求融入 AI 工具迭代，深度赋能 AI 影视工业化。

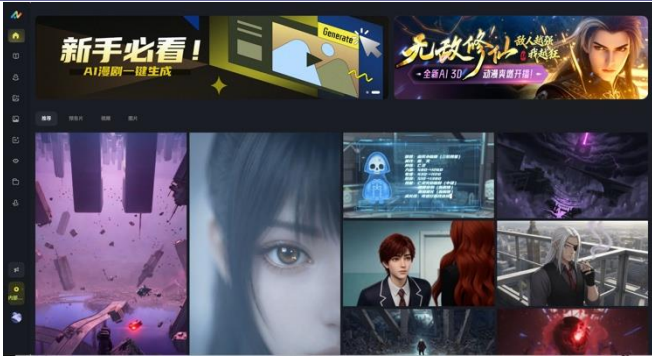
1) 中文在线：次元神笔赋能 AI 短剧工业化，海内外爆款作品持续落地

AI 视频工具：中文在线推出全栈 AI 内容创作解决方案次元神笔，该平台针对 AI 漫剧、真人短剧的内容创作 workflow，提供涵盖 IP 选品、分镜生成、角色建模、视频生成及智能配音的全流程 AI 功能。此外，自训练了针对北美市场的美漫风格 LoRA 模型，解决了国产漫剧出海的视觉本土化难题。根据公司内部测试和实践，公司开发的 AI 工具链整体可实现内容制作约 50%-70% 的时间节省、超 60% 的成本节省。

AI 内容：目前公司已运用 AI 技术创作累计超 250 部 AI 漫剧，2025 年公司 AI 漫剧作品全网累计播放量突破 60 亿。其中，《仙尊归来：只手镇人间》：公司首部 AI 3D 动漫，由次元神笔全程赋能制作，2025 年国庆档上线即登顶抖音动态漫热榜首位，

全网播放量突破 1 亿。《愤怒的吸血鬼》：专为北美市场打造的 AI 美漫风格漫剧，TikTok 上线 3 天播放量达 2.3 亿，登顶北美短剧热榜。

图表 58：次元神笔工作台



资料来源：中文在线，太平洋证券研究院

图表 59：AI 3D 动漫《仙尊归来：只手镇人间》



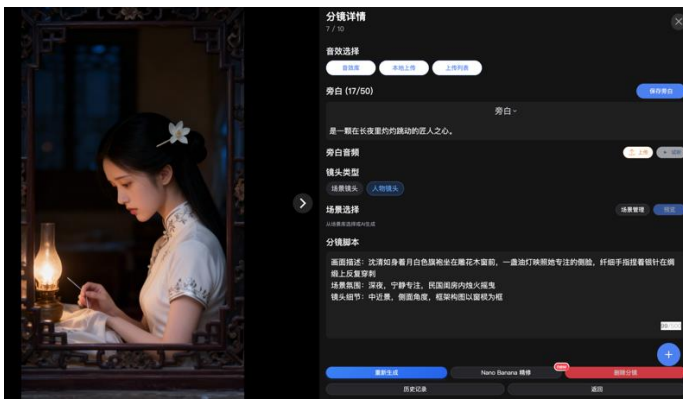
资料来源：中文在线，太平洋证券研究院

2) 博纳影业：博乐深度融合影视制作专业能力，AI 原生动画电影持续推进

AI 视频工具：博纳影业推出一键 AI 短剧智能生成平台博乐，该平台依托博纳 AI 自研算法、接入外部头部视频模型，深度融合生成式 AI 技术与影视制作专业知识，实现一键生成 AI 剧集。得益于公司在影视制作中积累的深厚经验，博乐结合创作者在使用过程中出现的真实需求，推出了：角色库及场景库功能，可将剧集中的人物角色及场景一键保存，实现跨集数的永久复用，并保证了角色特征与场景底色的一致。博乐成片预演系统，将成片修改环节提前到调用模型生成之前，提高了生成效果及效率。

AI 内容：公司成立全资子公司博越星纪蓝图，实现 AI 业务的市场化独立运作，目前的内容布局有：1) 2024 年 7 月公司推出国内首部全 AIGC 生成式科幻短剧集《三星堆：未来启示录》，全网播放量超 1.6 亿次，第二季已制作完成；2) 由博乐生成 AI 短剧，如《来自太平洋的你》、《关灯请闭眼》等，已在公司微信视频号、抖音官方账号等播出；3) AI 原生动画电影《三星堆：未来往事》于 2025 年 9 月 4 日在国家电影局立项公示，目前已进入制作收尾阶段，审核工作在同步开展。

图表 60：博乐工作台



资料来源：博纳影业 AIGMS 制作中心，太平洋证券研究院

图表 61：博乐创作作品部分海报展示



资料来源：博纳影业 AIGMS 制作中心，太平洋证券研究院

3.3 第三方工具公司：初创公司为主，聚焦垂类应用场景

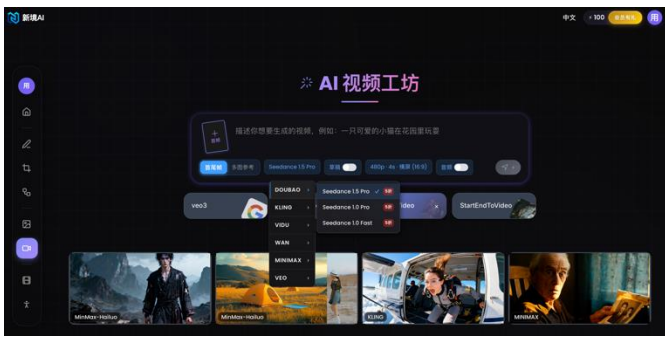
第三方工具公司以初创企业为主，专注短剧、漫剧等垂类场景，提供从剧本、分镜到成片的一站式创作工具。相较于视频模型、IP 及影视公司，其兼具 AI 工具开发的技术实力与垂类内容制作的实操经验，推动自研全流程工具在短剧、漫剧等细分赛道快速渗透。

1) 智灵新境：Agent 化实现 10 倍效率提升，优质作品陆续上线

AI 视频工具：智灵新境推出一站式视频短剧创作平台新境 AI，该平台外接 Veo、Seedance、可灵等主流 AI 多模态模型，开发了适用于短片创作的数十个实用工具，如分镜联想、影视光照优化等，形成覆盖全流程、支持精细化控制与优化的 AI 创作能力。新境 AI 以 **Agent 化的自动化处理为核心**，打造了多个短片 Agent，如剧本二创、剧本拆分、分镜提取等，提升内容创作效率。根据公司数据，相较于传统方式，智灵新境通过 AI 智能体实现了 10 倍效率提升，即传统+AI 视频工具的方式每日仅能完成 1 分钟内容的产出，且需专注于单一项目的制作。AI 智能体每日可完成 10 分钟以上内容的产出，且可并行操作多个项目。

AI 内容：新境 AI 已赋能独立创作者、漫剧团队、影视公司等各类客户，代表作品有：1) 《浮光》：AI 科幻短剧，凭借精致的 CG 画风、鲜明的风格及宏大的世界观，全网播放量已破 500 万；2) 《谁让他当鬼差的？》：S+级的 AI 玄幻穿越漫剧，红果短剧首日播放量近 400 万等。

图表 62：新境 AI 工作台



资料来源：新境 AI，太平洋证券研究院

图表 63：AI 科幻短剧《浮光》



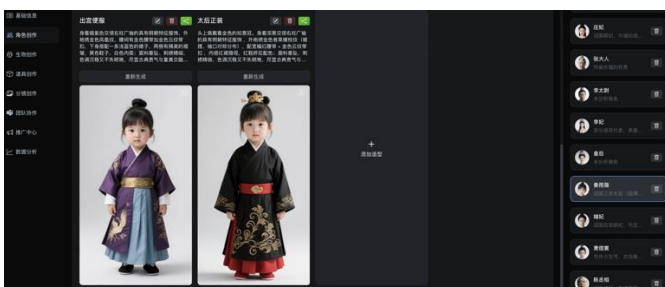
资料来源：《浮光》，太平洋证券研究院

2) 可梦智能：构建企业级 AI 短剧创作平台，AI 仿真人剧标杆作品频出

AI 视频工具：可梦智能推出专为企业级用户打造的 AI 辅助创作平台可梦 AI，该平台聚焦短剧、小说可视化两大核心场景，推出短剧创作模块和小说推文模块，通过 AI 技术打通从剧本、小说文本，到视频、可视化片段等成品内容的全流程创作。此外，平台具备通用宣发能力，通过为两大模块配套推广中心，实现描述需求即可生成专属海报、主题曲及歌词，无需依赖多工具切换。根据公司数据，运用可梦 AI，可实现 10 人团队 10 天完成 100 分钟短剧的高效产出。

AI 内容：可梦 AI 已赋能短剧公司、工作室等各类客户，打造出多部 AI 仿真人剧的标杆作品，如《斩仙台》AI 真人版：AI 神话仙侠短剧，上线仅 6 天抖音播放量破亿，是目前播放量破亿最快的 AI 仿真人剧。《我靠唱歌打脸全团》：AI 年代歌舞短剧，播放量突破 1.17 亿，ROI 超 1.1 等。

图表 64：可梦 AI 工作台



资料来源：可梦 AI，太平洋证券研究院

图表 65：可梦 AI 精品 AI 短剧



资料来源：可梦 AI，太平洋证券研究院

3) 灵境万维：自研 AI 原生动漫工业基座，构筑 AI 转绘漫剧优势

AI 视频工具：灵境万维推出一站式 AI 生产平台灵境 AI，该平台基于自研 AIGC 制作工具，提供视频创作、剧本创作、灵感选题三大功能，实现从选题、剧本创作到分镜成片的 AI 漫剧、AI 真人剧制作全流程。得益于自研 AI 原生动漫工业基座，以及制作中积累的大量优质动漫生产数据，公司在 AI 动漫领域具备领先优势。根据公司数据，相较于传统动漫制作，灵境 AI 可实现提效约 20 倍，降本约 90%，截至 2025 年 6 月，灵境 AI 动漫的每分钟制作成本仅数百元。

AI 内容：基于对目前短剧存量较大，AI 转绘漫剧对导演和编剧的依赖度较低、可实现大规模量产这两点判断，灵境 AI 开辟了 AI 转绘漫剧的业务线，打造出多部 AI 转绘漫剧的标杆作品，如《**武极天尊**》：改编自爆款武侠真人短剧，为行业早期案例。《**抢亲局中局**》：改编自爆款古言类真人短剧，播放量超千万等。

图表 66：灵境 AI 工作台



资料来源：灵境 AI，太平洋证券研究院

图表 67：AI 转绘漫剧《武极天尊》转绘效果



资料来源：灵境 AI，太平洋证券研究院

4. 风险提示

- 1) AI 技术发展不及预期的风险：**不排除 AI 视频模型的生成效果、一致性、物理规律模拟等发展放缓，导致 AI 应用渗透率提升不及预期的风险；
- 2) 内容监管趋严的风险：**不排除 AI 生成内容监管趋严、内容审核标准收紧，导致内容商业化落地进展不及预期的风险；
- 3) 行业竞争加剧的风险：**不排除 AI 技术发展加剧行业竞争程度，从而使得个别公司市场份额流失的风险。

投资评级说明

1、行业评级

看好：预计未来 6 个月内，行业整体回报高于标普 500 指数 5%以上；

中性：预计未来 6 个月内，行业整体回报介于标普 500 指数-5%与 5%之间；

看淡：预计未来 6 个月内，行业整体回报低于标普 500 指数 5%以下。

2、公司评级

买入：预计未来 6 个月内，个股相对标普 500 指数涨幅在 15%以上；

增持：预计未来 6 个月内，个股相对标普 500 指数涨幅介于 5%与 15%之间；

持有：预计未来 6 个月内，个股相对标普 500 指数涨幅介于-5%与 5%之间；

减持：预计未来 6 个月内，个股相对标普 500 指数涨幅介于-5%与-15%之间；

卖出：预计未来 6 个月内，个股相对标普 500 指数涨幅低于-15%以下。

太平洋证券股份有限公司

云南省昆明市盘龙区北京路 926 号同德广场写字楼 31 楼



研究院

中国北京 100044

北京市西城区北展北街九号

华远·企业号D座

投诉电话： 95397

投诉邮箱： kefu@tpyzq.com

免责声明

太平洋证券股份有限公司（以下简称“我公司”或“太平洋证券”）具备中国证券监督管理委员会核准的证券投资咨询业务资格。

本报告仅向与太平洋证券签署服务协议的签约客户发布，为太平洋证券签约客户的专属研究产品，若您并非太平洋证券签约客户，请取消接收、订阅或使用本报告中的任何信息；太平洋证券不会因接收人收到、阅读或关注媒体推送本报告中的内容而视其为太平洋证券的客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何机构和个人的投资建议，投资者应自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

本报告信息均来源于公开资料，我公司对这些信息的准确性和完整性不作任何保证。负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映研究人员的个人观点。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价或询价。我公司及其雇员对使用本报告及其内容所引发的任何直接或间接损失概不负责。我公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。本报告版权归太平洋证券股份有限公司所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登。任何人使用本报告，视为同意以上声明。